AD_____

Award Number: DAMD17-99-1-9294

TITLE: Automated Spot Mammography for Improved Imaging of Dense Breasts

PRINCIPAL INVESTIGATOR: Mitchell M. Goodsitt, Ph.D.

CONTRACTING ORGANIZATION: University of Michigan
Ann Arbor, MI 48109-1274

REPORT DATE: October 2004

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	October 2004	Final (15 Sep 1999 - 14 Sep 2004)

4. TITLE AND SUBTITLE
Automated Spot Mammography for Improved Imaging of Dense Breasts

5. FUNDING NUMBERS
DAMD17-99-1-9294

6. AUTHOR(S)
Mitchell M. Goodsitt, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Michigan
Ann Arbor, MI 48109-1274

E-Mail: goodsitt@umich.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

An automated spot mammography technique was developed to improve imaging of lesions within dense breast tissue. The initial goal was to build a device that would automatically detect suspicious regions in mammograms and immediately perform spot compression and collimation at that region to produce a second image that better distinguishes masses from overlapping tissues. Preliminary studies with a prototype device and breast simulating test objects showed promise, but spot compression didn't always separate the tissues as much as desired. We conceived an improved method that employs 3D stereoscopic digital spot mammography acquisition and display for superior perception and characterization of masses. Automated detection of suspicious regions was evaluated by comparing regions in patients' mammograms selected by radiologists with those selected by a computer program. There was good agreement between the radiologists and the computer A spot collimator to restrict the x-ray beam to the suspicious region for stereo spot image acquisition was designed, built, tested and refined. Finally preliminary studies were performed on a further evolution of the technique – utilizing tomosynthesis to reconstruct slices of the breast, with the eventual goal of developing automated spot tomosynthesis for enhanced sensitivity and specificity in breast cancer screening.

14. SUBJECT TERMS
Breast Cancer

15. NUMBER OF PAGES
76

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

(3) TABLE OF CONTENTS

(4) INTRODUCTION

The purpose of our project was to develop a novel technique for improved imaging of suspicious dense regions within full-field digital mammograms. The method we first proposed was automated spot compression mammography. The basic idea was to use a computer program to automatically detect any large dense region within a whole breast digital mammogram, and without moving the patient, take a second digital mammogram of only that region using automated spot collimation and automated spot compression. Preliminary studies of this imaging method using prototype devices and compressible breast simulating test objects showed promise, but the spot compression did not always result in the desired separation of masses from surrounding dense tissue. This prompted us to invent an alternative method that involved 3-D stereoscopic imaging and display with its inherent improved depth perception. The method, termed automated stereo spot digital mammography, produces better tissue position discrimination without spot compression. Thus, it does not require the use of a breast restraining device for the changeover from the large compression paddle to the smaller spot compression paddle which was requisite for the first method. Furthermore, since it does not employ a spot compression paddle, there is no need for automated positioning of such a paddle. Therefore, this alternative method is easier to implement. The stereo spot imaging procedure begins, as in the first method, with computer aided detection of a large suspicious region in the original digital mammogram. This is then immediately followed by the acquisition of a second 3D stereo digital mammogram of only that region using a more penetrating x-ray beam, automated spot collimation, manual x-ray tube shift for acquiring the left- and right-eye stereo images, and stereoscopic viewing of the images. Since only the dense region of the breast is exposed for the stereo spot mammogram, detector saturation in adjacent regions is not an issue. Restriction of the x-ray beam to a small "spot" area limits the amount of breast tissue exposed to additional radiation and decreases the volume of tissue that scatters x-rays thereby improving the contrast between a lesion and its surrounding dense breast tissue. Stereoscopic image acquisition and display enables the radiologists to view the suspicious regions in 3D. This reduces the tissue superposition problem inherent in conventional single projection mammography, wherein lesions can sometimes be camouflaged by overlying and/or underlying tissues, and normal tissues can sometimes overlap producing an object in the image with a lesion-like appearance. Furthermore, in comparison with conventional spot compression, the automated technique should produce more accurate spot imaging of suspicious regions because it obviates the repositioning and recompression of the breast for the spot images, and the "spot" location is determined by computer analysis of the digital full-breast image rather than estimated by the technologist by

eye from a radiograph. In brief, the project entailed (1) development of software to analyze digital mammograms and automatically delineate suspicious dense tissue regions, (2) development of hardware to automatically collimate the x-ray beam to those regions while maintaining the breast in the same position and compression as for the original full-breast digital mammogram, (3) development of a stereo spot technique for imaging the suspicious regions, and (4) development of display software on a stereo workstation for viewing the images of the suspicious regions in 3D.

(5) BODY

Following is a summary of our accomplishments for each of the tasks outlined in our Modified Statement of Work (dated January 7, 2002). Since this was an evolving methodology that involved major changes in equipment and experimental design, the project took an additional 2 years to complete (5 years instead of 3). We had applied for and received two 1 year no-cost time extensions.

A) Task 1: Develop software to recognize and delineate dense breast regions in full-breast mammograms.

The automatic detection of suspicious regions is a crucial first step of the automated stereo spot imaging technique. In order to evaluate the accuracy of the detection of such regions with a modified in-house developed computer aided detection (CAD) program, we needed to perform a study comparing regions selected by the computer program with those selected by a group of expert radiologists. During the first 4-years of this project, we developed computer software that permitted the radiologists to use a computer mouse to trace up to 3 suspicious regions in each mammogram, and a companion program to analyze the agreement between the regions selected by the radiologists with each other and with the computer and also with true regions of interest that were traced by a radiologist based on film mammograms, biopsy images and pathology results. The program development and results are described in detail in the annual reports for these years and in a publication in the Medical Physics Journal ("An Observer Study Comparing Spot Imaging Regions Selected by Radiologists and a Computer for an Automated Stereo Spot Mammography Technique." Medical Physics 2004; 31: 1558-1567).[1] Some of the important points of the study and features of the computer programs are described below:

i) TRACEIMAGE, program for radiologist tracing of boundaries of suspicious regions.

As shown in Fig. 1, the TRACEIMAGE program that we developed allowed radiologists to trace up to 3 regions for spot imaging in each mammogram. These regions are displayed in red on the mammogram. After completing the traces, the radiologist can select any one of the regions for modification. In the case illustrated in Fig. 1, area 1 is selected. The trace color for the selected area changes color to a bright red. The radiologist can alter the shape of the selected region by using the cursor (yellow arrow in mammogram) to grab an anchor point on the original trace and move it to a new location.

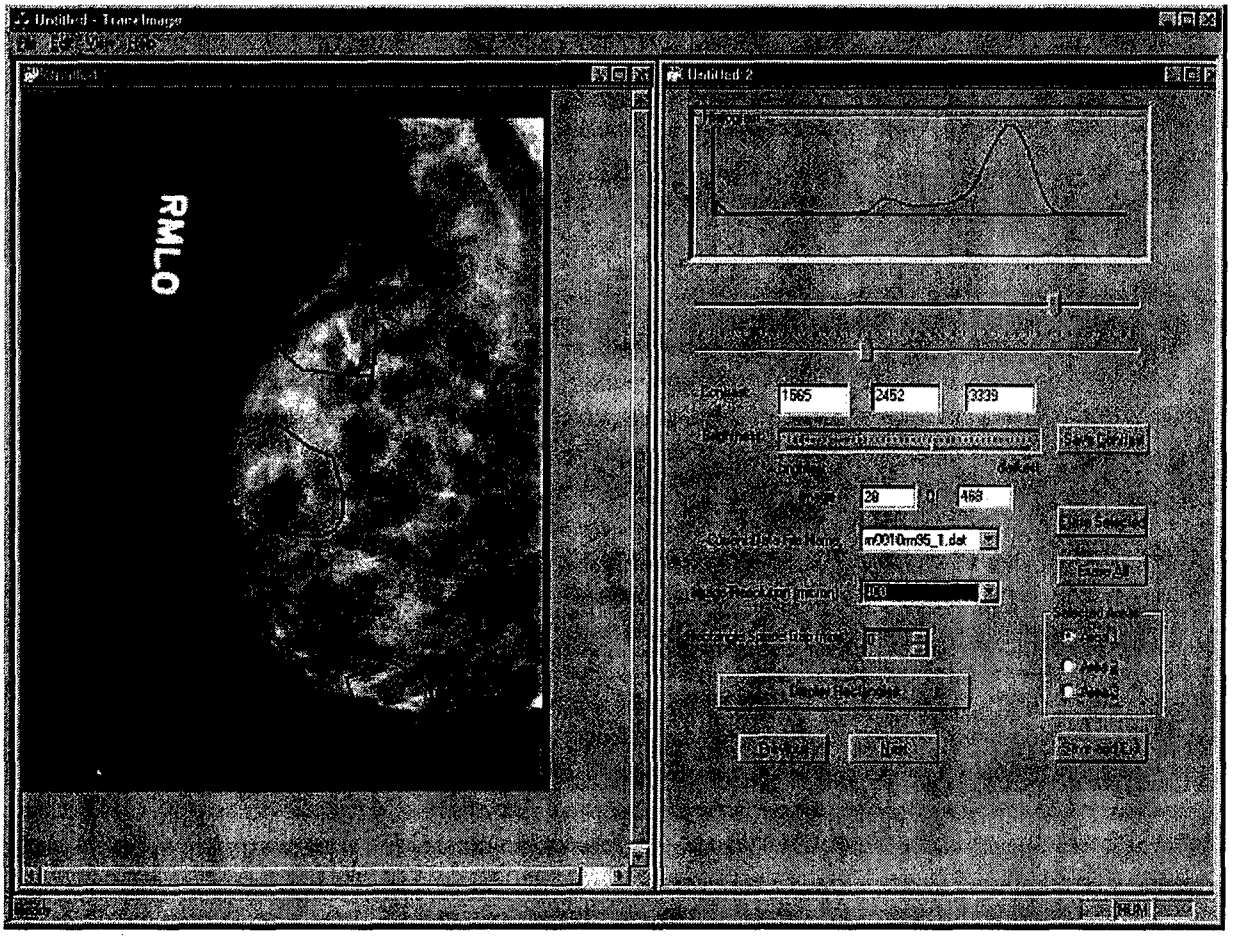


FIG. 1. Graphical User Interface (GUI) display showing one of the digitized mammograms in this study.

A histogram of the pixel values within the breast region of the image is displayed (top right) and sliders (beneath histogram) are employed for adjusting the range of pixel amplitude values that are mapped into an 8-bit output for display. That is, the sliders permit the radiologist to adjust the brightness and contrast of the images that they analyze. There is also a pull-down menu for selecting the image resolution (the resolution displayed in Fig. 1 is 400 microns.) Extensive initial studies using images with 800 micron pixel size showed that radiologists preferred the option to view higher resolution to make their selection of suspicious regions. We therefore modified the program to provide the radiologists with the option to trace the outlines in images displayed with pixel sizes of 200, 400 or 800 microns. Since the entire image could not be displayed on the monitor at 200 micron resolution, we developed horizontal and vertical sliders for panning the image. Examples of higher resolution (200 micron and 400 micron pixel size) images are shown in Fig. 2, below.

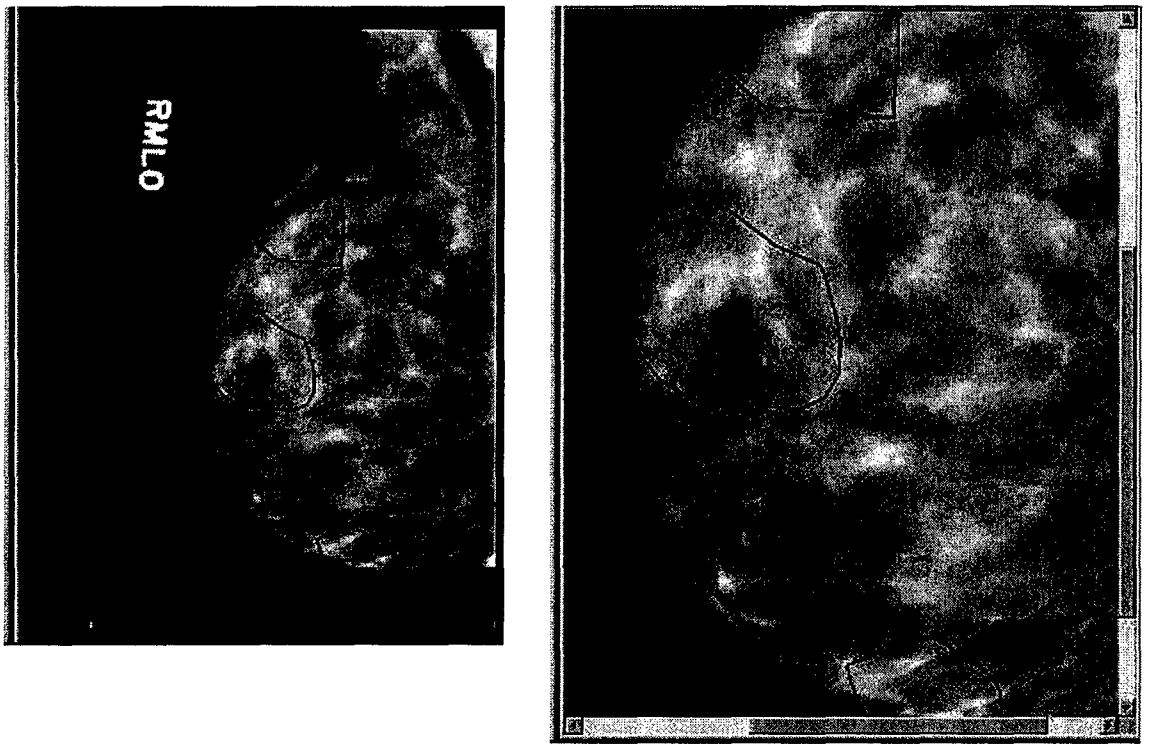


FIG. 2. Example of GUI display of the same image at 400 micron (left) and 200 micron (right) resolution. Radiologists could use either display resolution for each image in the study and could switch between the resolutions during their analysis.

The TRACEIMAGE computer program was very user friendly and was well-liked by all 5 radiologists who participated in our study that was published in Medical Physics.

ii) Image Data Set

The images for our studies were digitized film mammograms that have been employed previously in the development of our group's computer aided detection (CAD) system. We obtained Institutional Review Board approval to employ these images in our studies. The film set included craniocaudal (CC) view and the mediolateral oblique (MLO) view mammograms of both breasts of patients at our clinics. The mammograms were digitized with a LUMISYS 85 laser film scanner (Lumisys, Inc., Sunnyvale, CA) at a pixel size of 50 μ x 50 μ. The pixels were binned to generate the 200, 400 and 800 micron resolution images that were analyzed by the radiologists. To keep the reading time reasonable (i.e., about 3 hours), we had the observers perform the study on 200 mammograms. All of the observers also repeated the study 3 to 10 months later for an evaluation of their reproducibility.

iii) Computer selection of suspicious regions

The computer-selected suspicious regions were determined with a mass detection computer aided detection (CAD) program developed at the University of Michigan.[2] This program consists of 4 steps. First, the digitized mammogram is processed with a density weighted contrast enhancement (DWCE) filter that adaptively enhances local area contrast in order to emphasize mammographic structures. Second, an edge detection algorithm is employed to define the borders of the enhanced structures, resulting in a set of detected structures. Third, a local refinement algorithm, which includes erosion and K-means clustering, is applied to the detected structures to improve the accuracy of the borders and to split large connected regions. Fourth, the refined detected objects are classified as masses or normal breast structures based on the input of extracted morphological and textural features into a linear discriminant classifier. Potential masses are identified using decision thresholds that are based on the linear discriminant classifier score and the maximum number of marks allowed per image. For the auto spot observer study, we adjusted the detection threshold of the CAD program [2] to mark between 0 and 3 regions in each image, with 3 the most likely number.

iv) True regions of interest

Biopsy proven true masses (both benign and malignant) were identified in the digitized mammograms of our study by a radiologist from analyses of the mammograms along with associated pathology data and biopsy images. Rectangular regions of interest encompassing the masses were determined for analysis and display. These regions are referred to as true regions of interest (TROIs). Of the 200 images that were evaluated by the radiologists in our observer study, 98 images contained TROIs. There were 83 images with a single true mass, 13 with two true masses, and 2 with three true masses. Thus, there was a total of 115 (= 83 + 13 x 2 + 2 x 3) TROIs. Out of the 98 images with TROIs, 51 (52%) contained malignant masses

v) DENSECOMP program for displaying and analyzing the agreement between the TROIs and the ROIs selected by the radiologists and the computer

A computer program named DENSECOMP was written to display the ROIs selected by the radiologists and the computer and also the TROIs. In addition this program analyzed the degree of agreement between the respective ROIs.

a) Display

The program was designed initially to display the ROIs selected by up to 3 readers at once using different colors. For example, the traces for one radiologist would be filled-in as red, those for a second radiologist would be filled-in as green, and those for a third radiologist would be filled-in as blue. Anywhere in the image where all three of the readers' traces intersected would be displayed as white, intersections between red and green would be displayed as yellow, and those between red and blue as magenta. As shown in Fig. 3a, below, the displayed intersections were sometimes complex and difficult to interpret. We added a feature to the DENSECOMP program which permits display of only the outer borders of the ROIs with color codes. (See Fig 3b). This made interpretation of the intersections easier, but they were often still complicated.

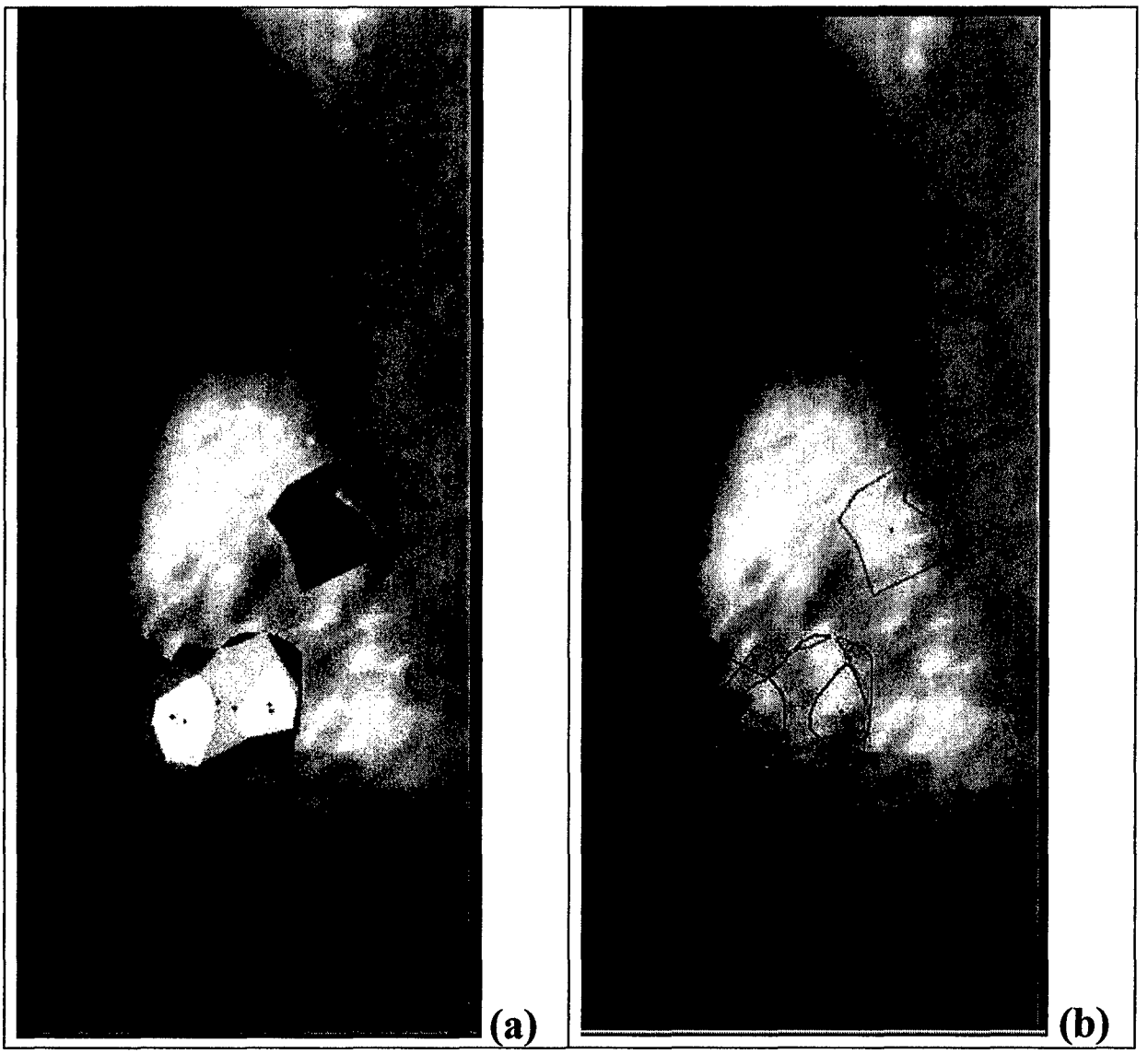
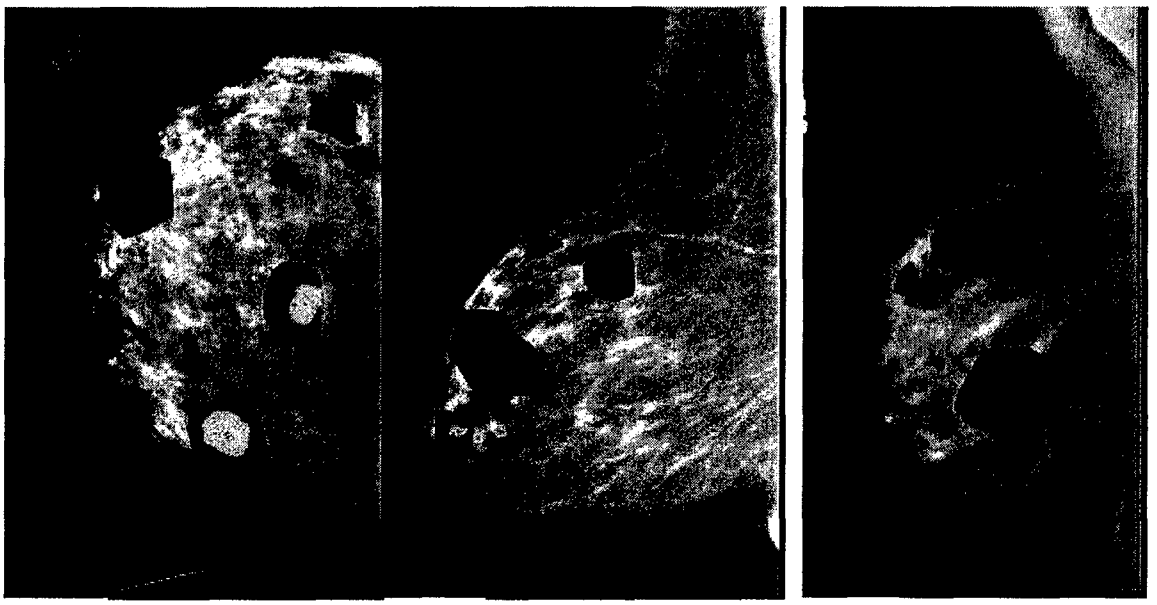


FIG. 3. Complex intersections of ROIs selected by 3 radiologists. In (a), ROIs for one radiologist are filled-in in red, those for the second radiologists are filled-in in green and those for the third radiologists are filled-in in blue. Intersections between all 3 are shown in white, those between red and green as yellow, those between blue and green as cyan, and those between red and blue as magenta. The DENSECOMP program also permits display of just the borders of the ROIs as shown in (b).

We therefore decided to display the ROIs for at most 2 "readers"- either for 2 radiologists, a radiologist and CAD, a radiologist and the TROI, or CAD and the TROI. Examples comparing the ROIs selected by a radiologist and the computer are shown in Fig. 4, and between the TROIs and ROIs selected by the 5 radiologists and between the TROIs and the ROIs selected by the computer are shown in Fig. 5.



(a) (b) (c)

FIG. 4. Examples of ROIs selected by a radiologist and the computer. The radiologist ROIs are filled-in in red, the computer (CAD) selected ROIs are filled-in in green, and the intersection areas are displayed in yellow. In (a) the radiologist and computer agreed on 2 of the 3 ROIs, in (b) they agreed on one ROI, and in (c) they disagreed on all 3 ROIs.

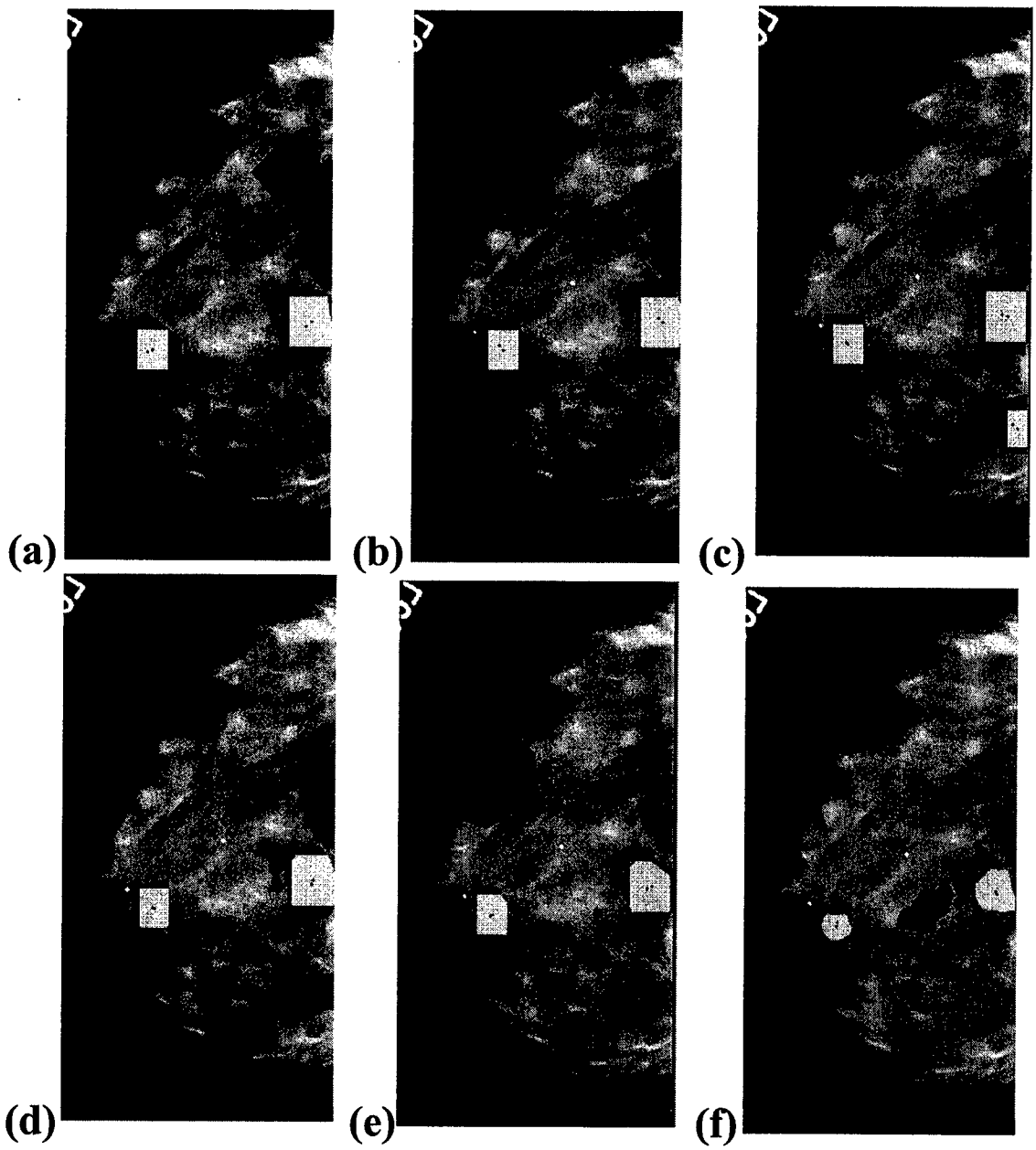


FIG. 5. Comparisons of ROIs selected by radiologists with true mass regions (a-e) and the CAD program with true mass region (f). The radiologist and CAD selected regions are filled-in in red, and the true regions are filled-in in green. The intersections are displayed in yellow. Note that for this particular mammogram, radiologist (d) chose to select 2 regions instead of 3. Also, radiologist (c) was the only one whose selected regions intersected all 3 TROIs.

b) Analysis

The DENSECOMP computer program computes overlap indices between the radiologist-selected ROIs and the CAD-selected ROIs. Furthermore, it computes overlap indices between the radiologists-selected ROIs and the TROIs and overlap indices between the CAD-selected ROIs and the TROIs. Initially, the overlap index that we used was the ratio of the areas of intersection between corresponding ROIs to the area of the union of those ROIs. We later decided to use a different index defined as the ratio of the area of the intersection between a spot region of interest selected by a radiologist and the spot ROI selected by the CAD program divided by the area of the CAD ROI. This index is thus 100% whenever the CAD region is completely contained within the radiologist selected region, representing complete agreement. Examples of 100% overlap are shown in Fig. 3a and b. We believe our new index more accurately depicts the amount of agreement between the ROIs than our previous intersection/union index. The reason is that even if the CAD ROI were completely contained within the radiologist's ROI, the previous index could be very small if the radiologist-selected region were fairly large. We also developed a new overlap index between the radiologist's or CAD ROI and the TROI. This index was defined as the area of intersection divided by the smaller area. This definition yields an overlap index of 100% if the radiologist's ROI or CAD ROI is completely contained within the TROI or if the TROI is completely contained within the CAD or radiologist's ROI. It is an indication of the degree to which there is a "hit" between the radiologist's or CAD's ROI and the TROI.

c) Observer study

Five MQSA qualified radiologists performed the observer study using our TRACEIMAGE software to trace up to 3 regions for spot imaging in 200 digitized mammograms. This study is thoroughly described in our Medical Physics paper [1], and the results are summarized here. Each radiologist performed the study twice with the reading session separated by a minimum of 3 months to minimize memory effects. This allowed us to determine the reproducibilities of the radiologists'' region selections. For the computation of the overlap indices between radiologist and CAD regions, all possible pairings of the selected regions were considered, and the indices were ordered from largest to smallest. For the radiologist vs. TROI and CAD vs. TROI cases, we computed the largest overlap indices individually for the first TROI, second TROI, and third TROI in the image

The average values of the largest overlap indices between the radiologist selected ROIs and the CAD selected ROIs are listed for each radiologist and each reading session in Table I, below.

Table I. Overlap indices between the ROIs selected by the radiologists and by the CAD computer program. The averages of the largest overlap indices for each image are listed for each radiologist at each reading session. These are averages for 200 images and up to 3 ROIs per image.

	Radiologist a	Radiologist b	Radiologist c	Radiologist d	Radiologist e
Reading #1	94%	84%	51%	45%	82%
Reading #2	98%	86%	47%	43%	65%

The agreement between the radiologist selected ROIs for spot imaging and the CAD selected ROIs had a wide range (43% to 98%). The primary reason for this is that the probability of overlap is highly dependent upon the numbers and sizes of the ROIs selected by the radiologist, with greater numbers and ROI sizes corresponding with a greater likelihood that there would be more overlap between the radiologist and CAD selected ROIs. Although each radiologist could select up to 3 ROIs in each image, we found there was a large range in the actual number of ROIs selected by the various radiologists. (The mean number of ROIs/image ranged from 0.66 for one of the radiologists to 2.94 for another.) In analyzing the data, we found there was a very high correlation of 0.99 between the overlap indices in Table I (above) and the average number of ROIs selected per image by the 5 radiologist observers.

The reproducibility of the overlap indices between the radiologist and CAD selected ROIs is very good (within 4% for 3 radiologists and 17% for one) indicating the radiologists are consistent in their selection of ROIs.

The average overlap indices between the TROIs and the ROIs selected by the radiologists and the CAD readers are listed in Table II, below.

TABLE II. Overlap indices between the ROIs selected by the readers and the True Regions of Interest.(TROIs). The averages of the largest overlap indices for each TROI in each image are listed, as well as the averages for all TROIs. The overlap index in this case is defined to be the area of the intersection divided by the smaller of the TROI and reader ROI areas, in percent. There were 83 images with one TROI, 13 with two TROIs and 2 with three TROIs.

	TROI #1 n=83	TROI #2 n=13	TROI #3 n=2	All TROI n=115
Radiologist a				
Reading #1	88%	58%	50%	84%
Reading #2	94%	74%	46%	90%
Radiologist b				
Reading #1	88%	70%	49%	85%
Reading #2	88%	66%	46%	84%
Radiologist c				
Reading #1	70%	42%	47%	66%
Reading #2	71%	44%	44%	67%
Radiologist d				
Reading #1	69%	45%	0%	65%
Reading #2	66%	48%	50%	64%
Radiologist e				
Reading #1	87%	50%	37%	81%
Reading #2	85%	67%	47%	82%
Computer (CAD)	78%	53%	0%	73%

The agreement between the radiologist selected ROIs and the true mass regions (TROIs) is good. For the entire set of 115 TROIs (Table II, column 4), the overall average overlap indices for the radiologists' ROI vs. TROI comparisons ranged from 64% to 90% with a mean value for all 5 radiologists of 76.8% +/- 10%, and the overall average overlap index for the CAD ROI vs TROI comparisons was 73%. The radiologists were very consistent in selecting similar regions at each reading session, with the overlap indices with all TROIs for four of the five radiologists being reproducible to within 1%. (Table II, column 4).

The good agreement between the CAD selected ROIs and the TROIs, which was comparable to the average performance of the 5 expert radiologists, indicates the CAD routines have promise in an automated implementation of stereo spot mammographic imaging.

B) Task 2: Develop Secondary Collimator

A secondary "spot" collimator was designed and built to be easily attached to and removed from our department's clinical GE Senographe 2000D digital mammography unit. A photograph of the collimator is shown in Fig. 6, below. It attaches to the face shield slot of the digital mammography unit (Fig. 7).

FIG. 6 Photograph of the spot collimator that was developed for this project. The blades for this collimator are driven with 4 stepper motors, which are also shown in the picture.

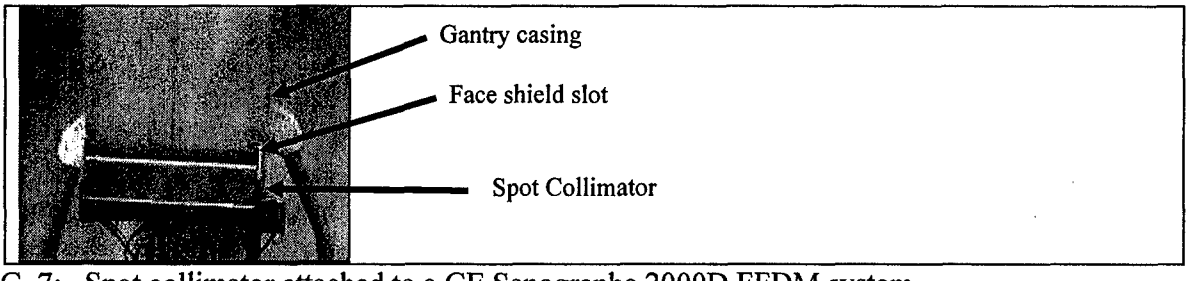


FIG. 7: Spot collimator attached to a GE Senographe 2000D FFDM system.

The collimator blades are positioned by an in-house developed computer program that controls the stepper motors that are attached to the collimator blades via worm and roller or belt mechanisms. The blade positions relative to the digital mammography image were calibrated by

using several digital mammography images that were acquired with the blades at known numbers of steps beyond their "home" positions when the collimator was physically attached to the GE full-field digital mammography system

Another computer program was written which displays a full-field digital mammography image, and allows the user (or CAD program) to position a rectangle in that image that represents the region of interest (ROI) for spot imaging. Furthermore, the program is interfaced to the stepper motor drives of the spot collimator and moves the blades to form the desired rectangular beam for spot imaging.

Because the x-ray tube is shifted to the left and right of the central ray for stereo image acquisition, it is necessary to analyze the imaging geometry to ensure that a mass located within the specified ROI is imaged entirely in the stereo mode. Fig. 8, below illustrates the projection geometry for stereo imaging.

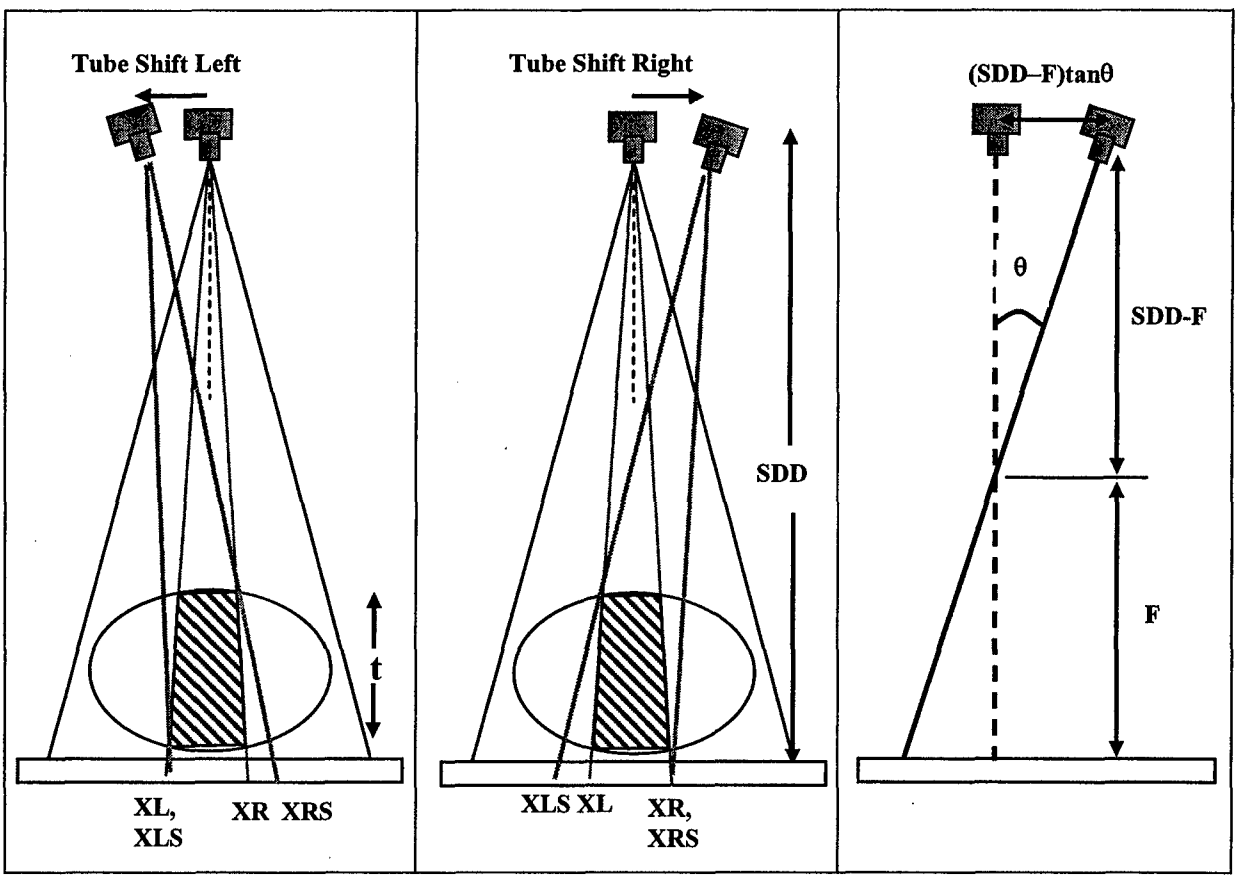


FIG. 8. Projection geometry for stereo spot imaging.

The cross-hatched region in this figure is the volume that projects to the image plane and forms the ROI on the standard (0 shift) mammogram. To guarantee that any object within this region is imaged during stereo imaging, the spot collimator blades must be positioned so that the

projections of their innermost edges correspond with the XLS and XRS image coordinates shown in the figure. We determined that the XLS and XRS coordinates for the left-shift image are given by the equations:

$$XLS = XL$$

$$XRS = XR + t * \tan\theta * \left(\frac{SDD - F}{SDD - t}\right)$$

The corresponding coordinates for the right-shift image are:

$$XLS = XL - t * \tan\theta * \left(\frac{SDD - F}{SDD - t}\right)$$

$$XRS = XR$$

where XL and XR are the transverse coordinates of the ROI for the 0° (non-stereo) image, SDD is the source-to-detector distance, θ is the stereo-shift angle, F is the fulcrum-to-detector distance, and t is the thickness of the breast. The Y-coordinates of the collimator blades and ROI in the chest wall-anterior direction do not change with the stereo shift.

We performed an experiment in which we took full-field and stereo spot collimated images of a custom-made stereoscopic breast phantom (CIRS, Inc.) containing overlapping masses. The experimental setup and full-field image with a marked ROI are shown in Fig. 9.

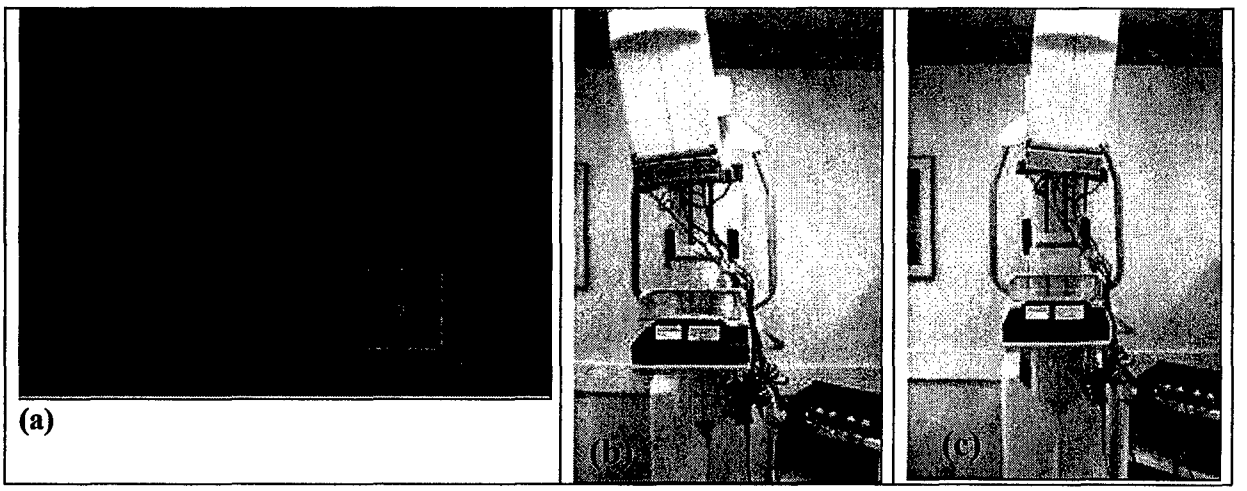


FIG. 9 (a) Full field digital mammogram showing user-selected rectangular ROI for spot imaging. Setup for left (b) and right (c) shift stereo spot acquisition.

The 0° spot, and 6° left-shift and 6° right-shift spot images of the ROI are shown in Fig. 10. The dashed lines in this figure indicate the borders of the original (non-stereo) ROI. Notice that the projected collimated image extends to the right beyond the original ROI for the left-shift spot image and extends beyond the original ROI to the left for the right-shift spot image, as desired.

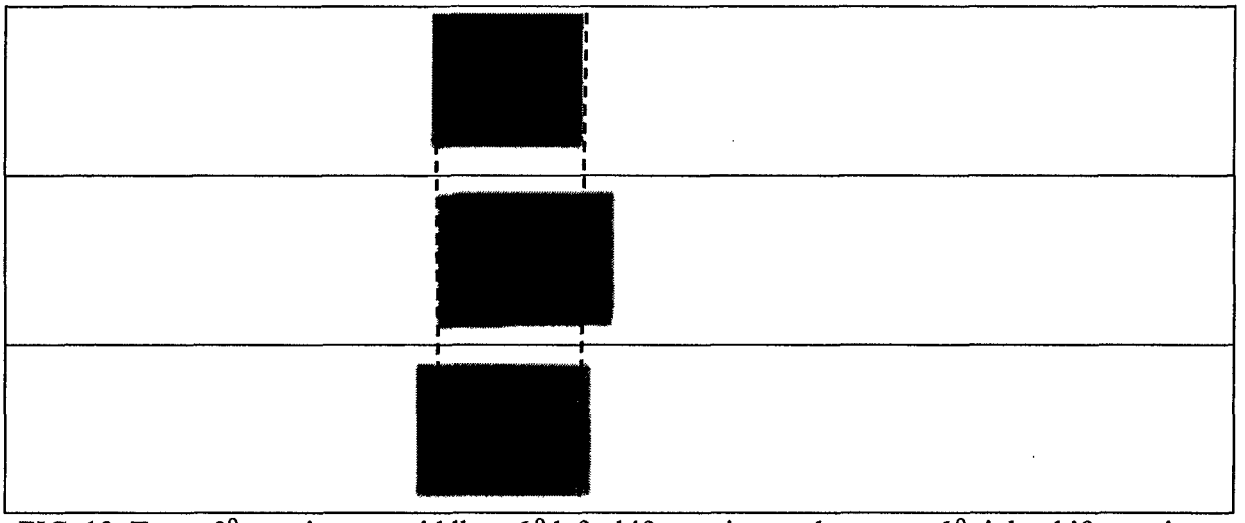


FIG. 10. Top = 0° spot image, middle = 6° left-shift spot image, bottom = 6° right-shift spot image

C) Task 3: Develop system to restrain breast during changeover from full-field to spot compression paddle and reduce tension as spot paddle presses against breast.

A prototype system for accomplishing the original task of restraining the breast during the changeover from full-field to spot compression paddle was designed and built during the first year of the project. This device and others that were built for our original automated spot compression concept are shown in Fig. 11. All of these devices were developed for use with a laboratory Fischer Biopsy system which was available for our research. The digital detector for this system was a computed radiography (CR) phosphor plate, which is not convenient for the procedure since the plate must be removed from the detector slot and transported across the hall to a readout device. The entire transport/readout process would take 60 seconds or more during which time the patient's breast would have to remain in compression. Towards the end of the first year of our project, a state-of-the-art GE full-field digital mammography system became commercially available, and one was installed in our breast imaging clinic. This system has a higher quality detector that is readout directly in about 10 seconds. Because of these superior characteristics for our application, we decided to switch to the GE system. We also conceived of a new method of spot imaging, automated stereo spot imaging. This new method did not require the use of a spot compression paddle and therefore did not require a breast restraining device, so there was no further development of the restraining device during the latter years of the project.

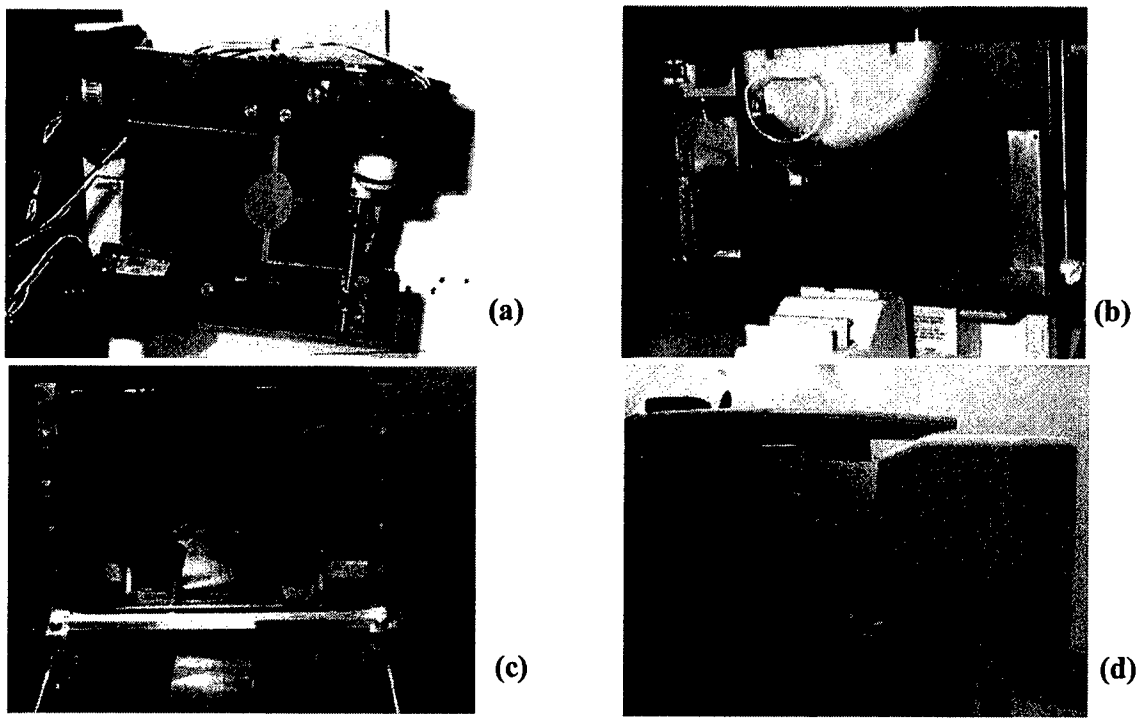


FIG. 11. First Generation Automated Spot Compression Instrumentation:
a) spot collimator with 2 stepper motor drives, one to rotate the collimator and the other to change the size of the collimator opening, b) x-y translator for manually positioning the spot compression paddle at the location determined by the CAD program, c) breast restraining device that keeps the breast in a fixed position during the changeover from the standard compression paddle to the spot compression paddle, d) spot compression device attached to our research Fischer breast biopsy unit which employs a CR plate as the digital detector.

D) Task 4: Develop system to position spot compression paddle

A prototype spot paddle positioner was designed and built during the first year of the project. It is shown in Fig. 11b, above. As with the restraining device, this spot paddle positioner was not needed for the new stereo spot technique that we pursued, so it was not developed further.

E) Task 5: Evaluate Automated Exposure Optimization (AOP) techniques for spot imaging application on commercial full-field digital mammography device.

The task of evaluating automated exposure optimization techniques in stereo spot mammography was not completed for two reasons. First we didn't receive a modular breast phantom from the manufacturer that came even close to meeting our design specifications until very late in the project. Second, when we did produce stereo spot images of this phantom, we found it difficult to impossible to distinguish the overlapping masses regardless of the dose (See Section F, below.) Thus the phantom made the imaging task too difficult to evaluate. We are working on

a different NIH funded project with GE Corporate Research to develop a combined x-ray tomosynthesis and 3-D ultrasound breast imaging device. We plan to image our modular phantom with this device to determine whether tomosynthesis is more successful than stereo mammography in distinguishing the overlapping masses.

Although we could not perform a comprehensive study to optimize the AOP technique for spot imaging, we performed an observer study comparing the ability of observers to distinguish positions of simulated fibrils at different radiation doses, stereo angles, and geometric magnifications. (Posters presented at the SPIE-Medical Imaging meeting at San Diego, CA in February, 2002) [3] and at the Era of Hope Meeting in 2002 [4]). We found that significant improvements in depth discrimination using a stereo shift angle of +/- 6 relative to +/-3 degrees, and geometric mag of 1.8X relative to contact (mag ~ 1X). Also, for large stereo shift angles and geometric magnification, the results were less dependent upon x-ray dose. These results indicate that the best stereo technique for the spot implementation would entail use of a stereo shift angle of +/- 6 degrees, geometric magnification of about 1.8X, and half the dose of a conventional mammogram for each image of the stereo pair (total dose = dose of a conventional mammogram).

F) Task 6: Develop breast phantoms

We designed a modular breast phantom that consisted of individual sections or slabs containing simulated masses and microcalcifications that could be positioned at different depths, and other sections containing simulated glandular tissue that would overlap the masses and microcalcifications. Another design criterion was that the phantom produce images with a patient-like mammographic appearance. All of the design requirements were very important for our work, because they would allow us to test the ability of stereo imaging to reduce the camouflaging effect of overlying and underlying tissues. We submitted our original designs for the breast phantom for stereo spot mammography research to the manufacturer (CIRS, Inc.) in November of 2002. The manufacturer found it difficult to meet our requirements and this resulted in an unexpected manufacturing delay of over 9 months. The phantom that we received did not meet our design goals. The glandular tissue regions were located too close to the outer edge of the phantom and had little or no overlap with the simulated masses, and the glandular tissues have a swirled appearance with sharp edges that is very different from the appearance in real mammograms. (See images in Fig. 12, below.)

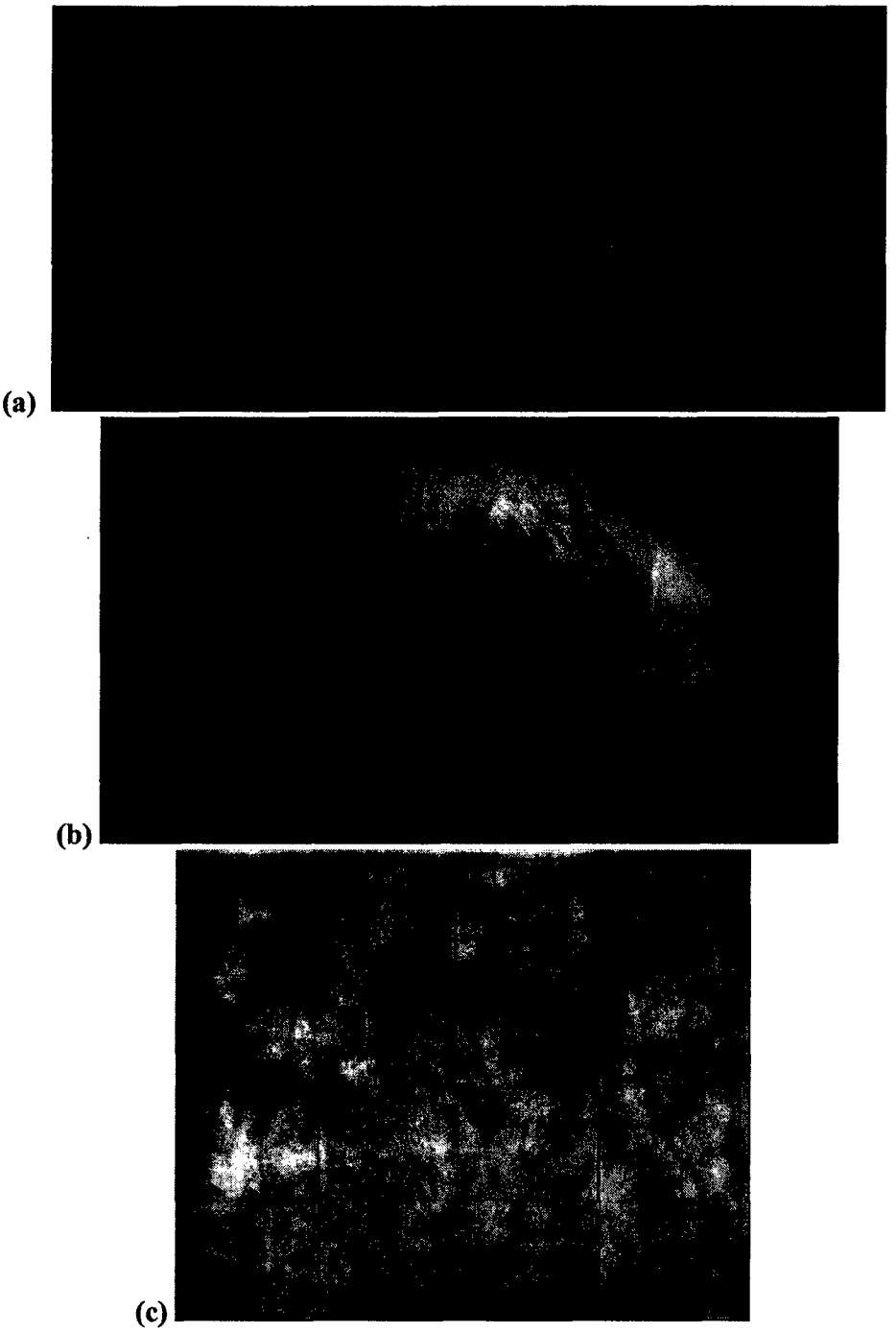


FIG. 12. (a) Desired mammogram appearance of phantom from our original design specifications, (b) Mammogram of the first phantom (c) X-ray image of a different phantom built by the same manufacturer for another client. Our original design specification called for an image texture similar to the texture in this image, which is also similar to that of the patient mammogram in (a). This mammogram (c) was published in reference 5. [5]

We contacted the manufacturer and requested that they make changes to improve the phantom. The manufacturer spent more than 1 year developing a revised phantom, which we received in March of 2004. We found the glandular regions of this new phantom to be so dense that they entirely obscured underlying or overlying masses. The manufacture agreed to try again, and we received new sections at the end of April of 2004. Images of two individual sections containing the masses (one of which is from the original phantom and was never revised by the manufacturer) and an image of the entire modular phantom with the slabs superimposed is shown in Fig. 13. As shown in part c) of this figure, the phantom has a more realistic mammographic appearance and the glandular regions are not too dense. However, the glandular regions do not sufficiently overlap the masses.

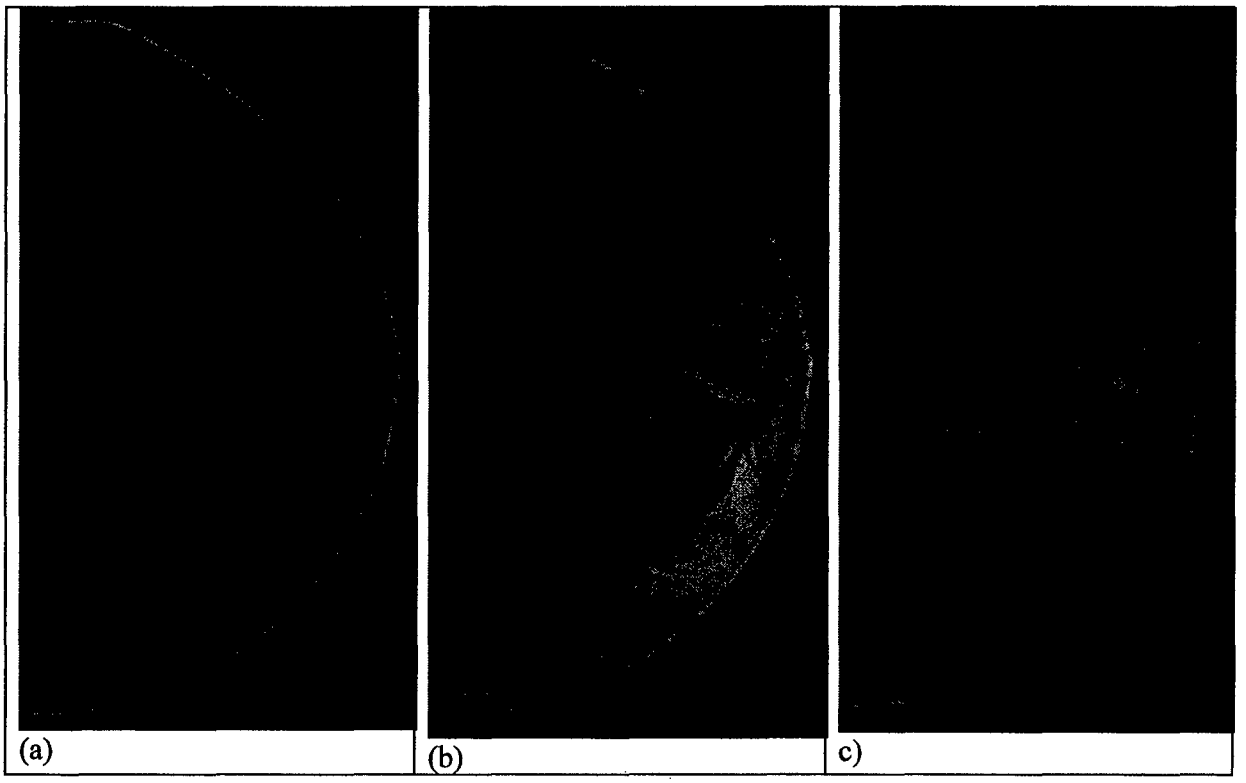


FIG. 13. (a) Digital mammogram of individual slab containing simulated spiculated masses, non-spiculated masses and microcalcifications, (b) digital mammogram of a different individual slab containing simulated masses and swirly dense glandular regions c) digital mammogram of our current modular breast phantom containing five 1-cm thick slabs including those in (a) and (b) separated by 2 cm.

Stereo spot images of one set of overlapping masses in the phantom with the full-field digial mammogram in Fig. 13c are shown in Fig. 14b and 14c.

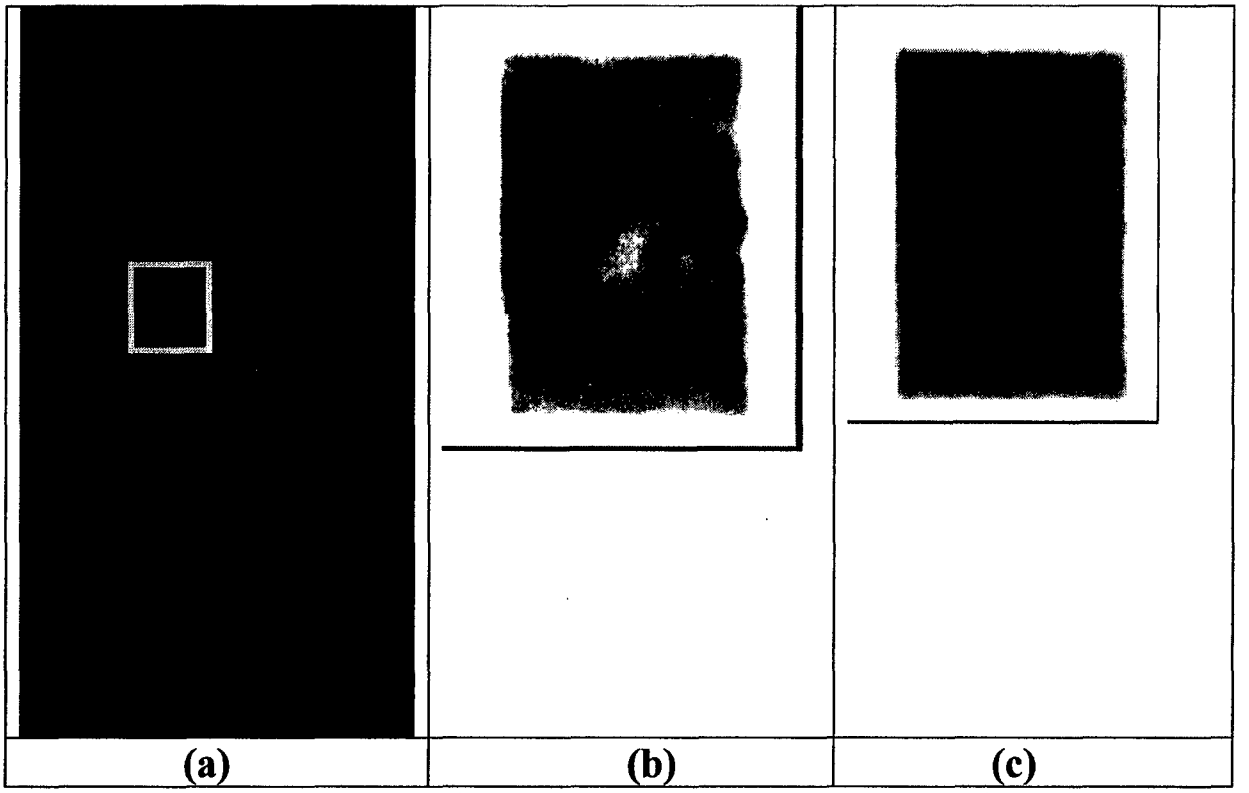


FIG. 14. a) Digital mammogram with overlapping masses of interest indicated by yellow box, b) stereo spot image of ROI corresponding to an x-ray tube shift of 6 degrees to the right, c) stereo spot image of ROI corresponding to an x-ray tube shift of 6 degrees to the left. (Left corresponds to up in this figure, and right corresponds to down.)

Notice how difficult it is to distinguish the overlapping masses in parts b and c of this image. One mass is easily seen, but the second is a very low contrast blur with poorly defined margins. Contrast this with the images in Fig 10 where two partially overlapping masses are easy to distinguish. Finally even when the right- and left-eye images in Fig 14 are viewed in stereo, the masses can not be distinguished.

G) Task 7: Explore possible advantages of using stereo-spot mammography instead of single-projection, spot compression with spot collimation mammography for imaging overlapping structures

The task of exploring possible advantages of using stereo-spot mammography instead of singe-projection spot compression and spot collimation mammography for imaging overlapping structures was completed during the second year of the project, and reported on in the second year annual report. We took full-field stereomammograms, spot compression and spot collimation stereomammograms and spot collimation with full breast compression stereo mammograms of a compressible breast phantom containing overlapping simulated masses. We employed 3D virtual cursors that we had developed to measure distances between the masses in the images and found that spot compression resulted in the masses being squeezed closer together rather than separating. This is also a possibility when spot compressing the human breast. On the other hand, spot and full-field stereomammography without spot compression did

not exhibit this problem. We concluded that spot compression may not be necessary or desirable, and pursued the 3D stereospot imaging alternative.

H) Task 8: Compare contact and magnification stereomammography

The task of comparing contact and magnification stereomammography was performed during the third year of the project. We compared the accuracies of depth measurements made with 3D stereo cursors in contact and magnification stereomammograms. We found that accuracy improved with geometric magnification and that using display magnification or zoom had no effect. We published a paper on our study entitled "The effects of stereo shift angle, geometric magnification and display zoom on depth measurements in digital stereomammography," (Medical Physics Volume 29 (11), November 2002, pages 2725-2734) [6] We also presented the results of our study comparing the ability of observers to distinguish positions of simulated fibrils at different radiation doses, stereo angles, and geometric magnifications in posters at the SPIE-Medical Imaging meeting at San Diego, CA in February, 2002) [3] and at the Era of Hope Meeting in 2002 [4]. We also performed an observer study comparing radiologist's abilities to distinguish the relative positions of simulated fibrils in contact and magnification stereomammograms and published a paper on this study entitled "Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study," Physics in Medicine and Biology 2003; 48(22): 3721-3734. [7]

I) Task 9: Develop and implement changes in automated spot collimator design for prototype spot stereo mammography image acquisition system.

For the task of developing and implementing changes in the automated spot collimator design for the prototype spot stereo mammography image acquisition system, we modified our computer program that was used to position the spot collimator blades to include the geometric relations described in section B. We obtained stereo spot images of user-defined regions of interest in 2 phantoms (See Figs. 10 and 14, above) and found the collimation system worked properly.

J) Task 10: Design second generation auto stereo spot mammography system

The original final task of this project was to design a second generation automated stereo spot mammography system. Based on the improved stereo perception and accuracy that we found with geometric magnification in our observer studies and 3D virtual cursor studies, the second generation automated spot mammography system should include the capability of stereo spot imaging the breast in a magnification mode. To satisfy this requirement as well as the requirement that the spot imaging occur within seconds of the original full-field digital mammogram while the breast is maintained in the same compression and position, we devised a design that would employ a movable C-arm with a fixed breast support plate. To illustrate how this system would work, consider a CC-view. After the original CC full-field digital mammogram is taken, a CAD program would determine a region of suspicion for spot imaging, the C-arm would then move down (the x-ray tube would move closer to the breast and the digital

detector would move further from the breast), creating a magnification geometry. If a grid were employed for the full-field digital mammogram, it would automatically be removed for the magnification spot image. The system would automatically shift the x-ray tube to the left for the left-eye image of the stereo spot mammography pair, and the collimator would automatically restrict the beam to the region of interest for that image. The left-eye digital spot mammogram would then be taken. Next, the x-ray tube would automatically shift to the right for the right-eye spot image, and the collimator would automatically restrict the beam to the desired region of interest for that image. Then the right-eye spot mammogram would be taken. The three images (contact full-field digital mammogram, left-eye spot mammogram, and right-eye spot mammogram) would be transferred to a workstation that displays both the conventional and the stereo images for review by the radiologist.

As described in section F, above, during the final 1-year no-cost time extension of this project, we investigated stereo imaging of a specially designed modular breast phantom. As illustrated in Fig. 14, we found it was very difficult to distinguish two overlapping masses in this phantom. Only one mass was clearly visible even in stereo. When we imaged the phantom with stereo magnification, there was no improvement. Also, we experienced significant eyestrain when viewing the magnification stereo images acquired with an x-ray tube shift angle of $\pm 6^{o}$. We had not experienced this problem in our previous stereomammography studies and believe this is due to the relatively thin (1.1 cm) phantom that was employed in those studies. [3,6,7] as compared with the ~5 cm thick modular breast phantom that was imaged. Most breasts are about 5 cm thick. Therefore, when imaging patients, a reasonable compromise between stereoacuity and eyestrain would be to employ a smaller shift angle of $\pm 3^{o}$.

Our inability to distinguish the overlapping masses in the modular phantom in both mag and non-mag stereomammography could be a result of the somewhat unrealistic pattern and structures of the breast phantom or that the stereo imaging technique is incapable of separating low contrast objects in close proximity. We plan to pursue this problem further when a better phantom becomes available in the future. In addition, we plan to investigate whether tomosynthesis works better than stereomammography in imaging this phantom. We will employ a tomosynthesis system that GE Corporate Research is developing for us in this investigation. That system should be installed by early spring of next year. During the final year of our project, we did perform a very preliminary study of tomosynthesis using a very simple shift-and-add reconstruction algorithm. Some of the results of that study are shown in the next section. The method appears very promising, and if it is shown to be superior to stereo imaging, we believe a future generation system should incorporate the capability of automatically performing contact and magnification automated spot tomosynthesis.

K) Task 11: Preliminary study of tomosynthesis

The final 1-year extension of the project afforded us the time to perform a preliminary study of tomosynthesis reconstruction. The GE Senographe 2000D digital mammography system that was available to us for this study is not ideal because it is not designed to perform tomosynthesis. In particular, the x-ray tube does not shift to a sufficient number of set (detented) angular positions and the collimator blades do not adjust to account for x-ray tube angulation. We solved the limited number of set angles by employing a digital protractor (Macklanburg-Duncan Model

Pro 360) to carefully set the x-ray tube to desired angles. We found this protractor to be accurate to about +/- 0.2 degrees. As shown in Fig. 15, below, the protractor was attached to the face shield of the GE gantry for our measurements.

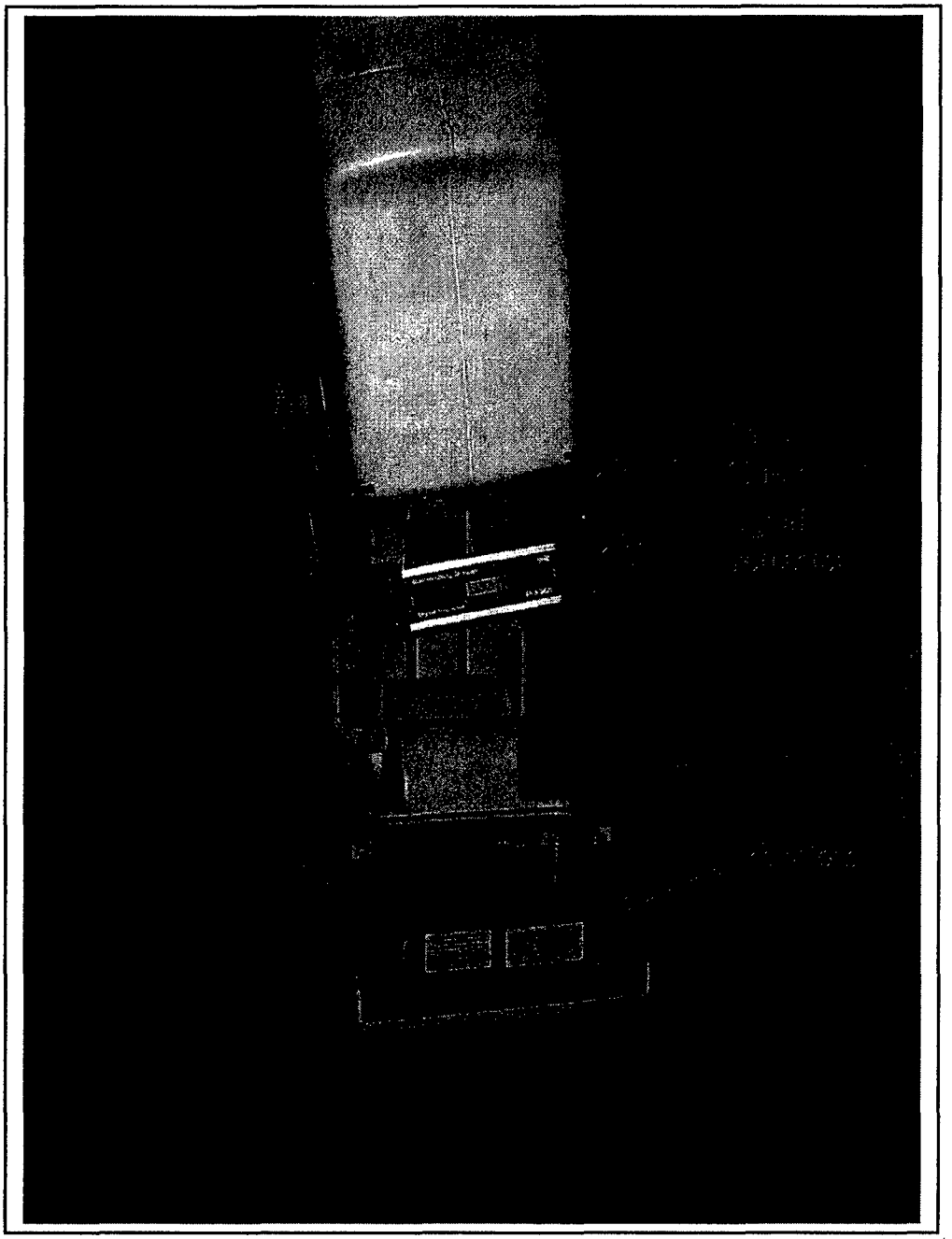


FIG. 15. GE digital mammography system with digital protractor attached to face shield and custom breast simulating phantom.

The phantom that we imaged consisted of simulated spiculated masses and microcalcifications at 4 different depths. We acquired 11 projections (images) at angles ranging from -15 degrees to +

15 degrees in steps of 3 degrees. A software program was written in Matlab to shift and add the images in order to reconstruct images at different depths. The manufacturer of the phantom sent us photographs showing the locations of the spiculated masses and microcalcifications in each section before the manufacturer poured a resin into the phantom mold to fill-in the spaces between the sections to create a solid phantom. Reconstructed tomosynthesis slices and the corresponding photographs are compared for 2 of these slices in Fig. 16, below. The agreement is very good. Masses in the particular slice are in-focus and others are out-of -focus, and we can also see the microcalcifications in the slices. This phantom has a uniform background. There fore, it is very easy to see blurs of out-of-plane objects in the reconstructed slices. More sophisticated reconstruction algorithms such as Maximum Likelihood algorithm used by Wu et al [8] and Matrix Inversion used by Dobbins [9] should reduce the artifacts and produce better images.

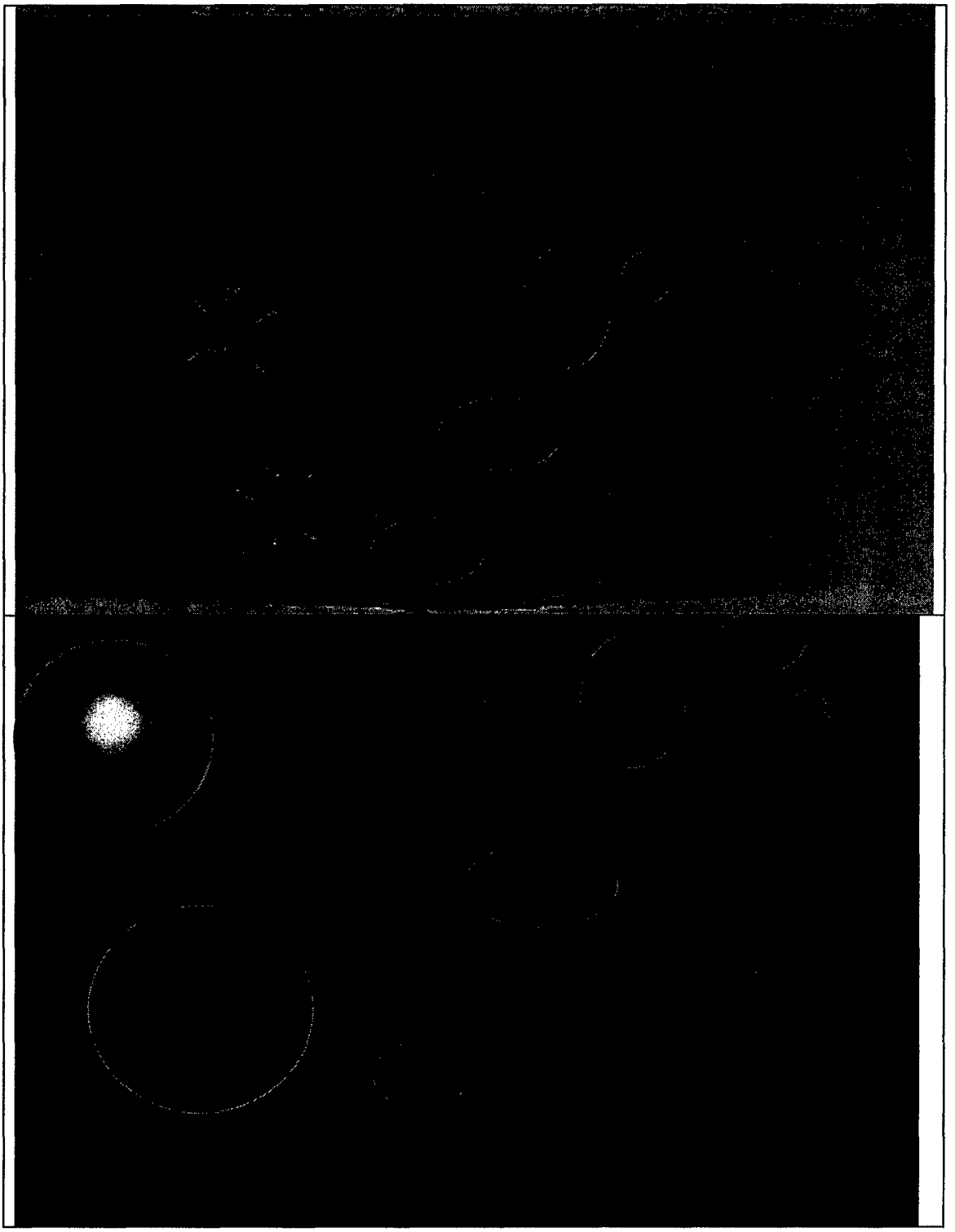


FIG. 16a. Photograph (above) and corresponding reconstructed tomosynthesis slice (below) of a breast phantom.. Microcalcifications are circled in green and spiculated masses in blue. Notice the spiculated masses at this depth are in-focus whereas spiculated masses at other depths are blurred.

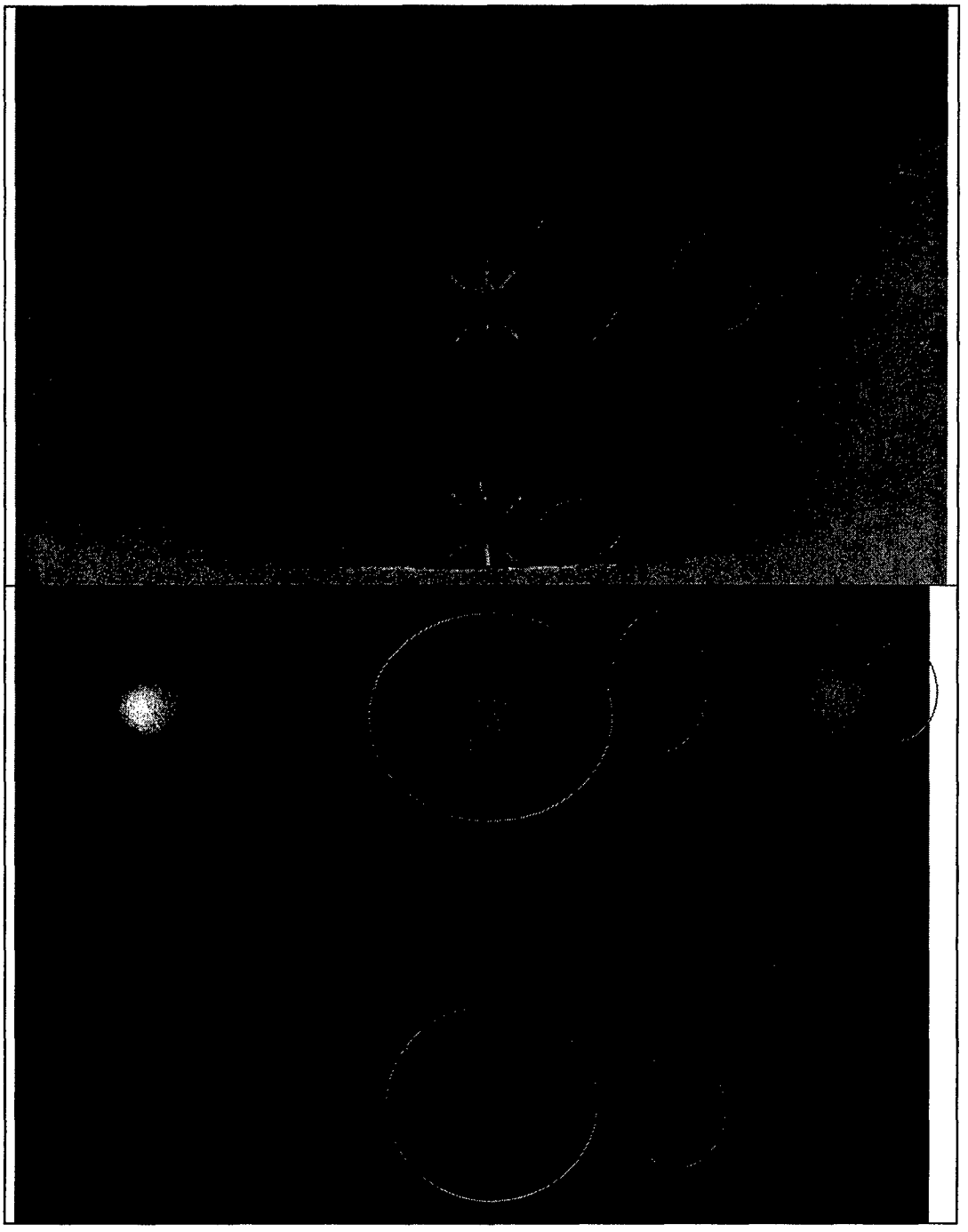


FIG. 16b. Photograph (above) and corresponding reconstructed tomosynthesis slice (below) of a different section in the phantom.. Microcalcifications are circled in green and spiculated masses in blue. Notice again that the spiculated masses at this depth are in-focus whereas spiculated masses at other depths are blurred.

(6) KEY RESEARCH ACCOMPLISHMENTS

- Developed prototype devices for implementation of automated spot compression and collimation on our research Fischer Biopsy x-ray system that we equipped with a computed radiography digital detector. The prototype devices included a spot collimator, a breast restrainer, and an x-y translation device for positioning a spot compression paddle.

- Modified research plans to develop automated spot compression on a state-of-the art GE full-field digital mammography (FFDM) system that was installed as a clinical unit in our radiology department towards the end of the first year of the project. The GE system produces images of superior quality and has a faster image readout time, which are advantageous for the eventual application of the method to patient imaging

- Performed stereoscopic spot compression, spot collimation and full-field compression experiments on the GE FFDM system with a compressible breast phantom and determined that spot compression sometimes fails to result in the desired separation of a mass from overlying or underlying tissue or masses. In fact, we found it can push the objects closer together making them less easy to distinguish. On the other had, 3-D stereoscopic imaging without spot compression should result in improved separation. This prompted us to conceive of a new spot compression technique, automated stereo spot compression which we pursued for the remainder of the project.

- Developed software for an observer study to compare suspicious regions for spot imaging selected by radiologists to those selected by a computer (computer aided detection (CAD)) and to true regions of interest corresponding with regions that were biopsied for which we had pathology results. The software included a program that enabled users to trace the borders of up to 3 suspicious dense regions in digitized mammograms, and a second program that displayed the regions selected by the radiologists, the computer (CAD) and the true regions of interest (TROI) in different colors on the digitized mammograms. This second program allowed us to visually compare the agreement between the regions selected by pairs of radiologists, each radiologist vs. CAD, each radiologist vs. the TROI, and CAD vs. the TROI.. The second program also computed overlap indices to quantify each of these comparisons.

- Performed an observer study in which 5 experienced breast radiologists used the above mentioned software to select suspicious dense regions in a set of 200 digitized mammograms. Quantified the results and found the overlap indices between the regions selected by the computer and the TROI were comparable to the average of the overlap indices between the radiologists and the TROI, indicating the computer determined regions could potentially be useful for a breast cancer screening technique that includes stereo spot mammography.

- Designed and had a machinist build an automated spot collimator that attaches to the GE digital mammography system.

- Wrote computer program to move spot collimator blades via a stepper motor controller. Calibrated the collimator blade positioner from images obtained with the spot collimator attached to the full field digital mammography unit. Wrote computer program to permit user to select a rectangular region in a full field digital mammogram for spot imaging.

- Investigated effects of stereo shift angle, geometric magnification and display zoom on depth measurements made in stereo mammograms using in-house developed 3D cursors. Found best results were obtained with stereo shift angle of +/- 6^{o} and geometric magnification of 1.8, and that display zoom was not beneficial.

- Compared observers' abilities to distinguish relative depths of simulated fibrils at different radiation doses, stereo angles, and geometric magnifications. Found significant improvements for stereo shift angle of +/- 6 vs +/-3 degrees, and geometric mag of 1.8X vs. contact (mag ~ 1). Also, for large stereo shift angle and geometric mag, results were less dependent upon x-ray dose.

- Based upon the above two stereomammography investigations, we concluded that the best stereo technique for the spot implementation would entail use of a stereo shift angle of +/- 6 degrees, geometric magnification of about 1.8X, and half the dose of a conventional mammogram for each image of the stereo pair (total dose = dose of conventional mammogram).

- Analyzed geometric considerations for stereo spot imaging to insure complete coverage of suspicious region in the pair of left and right-eye images of the stereo mammogram. Implemented the derived equations in software and performed tests to verify proper coverage. Found system worked well.

- Designed and had phantom manufacturer build a modular breast phantom containing simulated masses, microcalcifications and glandular regions for stereo spot imaging experiments. Despite the manufacturer 's best efforts they were unable to build us a phantom that met all of our design specifications. They have sent us 4 editions of the phantom. Some of the problems have included phantom sections with dense glandular regions that were located too close to the periphery and did not obscure the masses, dense regions that had a swirly appearance with very sharp edges which did not simulate the desired blurred patchy densities with soft edges in real mammograms, dense regions that were too radio-opaque, and masses placed in incorrect locations. These problems prevented us from completing some planned studies on optimum exposure control settings. However, as discussed above, with a different thin (1.1 cm thick) phantom, we found that, for large and small stereo shift angles and geometric magnification, stereo perception does not strongly dependent upon x-ray dose. With the present ~5 cm thick phantom, we did find that there was significant eyestrain when viewing stereo images acquired with geometric magnification and a large shift angle of $\pm 6^{o}$, so a smaller shift angle is preferred. Therefore, a compromise stereo spot technique could be the combination of a stereo shift angle of +/- 3 degrees, geometric magnification of about 1.8X, and half the dose of a conventional mammogram for each image of the stereo pair.

- Proposed spot tomosynthesis as an alternative to spot stereomammography for improved mass discrimination. Performed preliminary study of tomosynthesis with our GE FFDM system using digital protractor to set x-ray tube angles and simple shift-and-add tomosynthesis reconstruction. Results are promising and we plan to investigate the spot tomosynthesis method further in the future when our group obtains a research tomosynthesis system.

- Proposed a movable C-arm as a method for acquiring geometric magnification in a second generation automated stereo spot mammography system. Magnification is desired based on our studies that showed stereo conspicuity and accuracy improve significantly with geometric magnification.

(7) REPORTABLE OUTCOMES

Papers

a) Published in peer-reviewed journals

1) Zhou C, Chan HP, Petrick N, Helvie MA, Goodsitt MM, Sahiner B, Hadjiiski LM. Computerized image analysis: Estimation of breast density on mammograms. Medical Physics, 2001; 28: 1056-1069.
2) Goodsitt MM, Chan HP, Darner K and Hadjiiski LM. The Effects of Stereo Shift Angle, Geometric Magnification, and Display Zoom on Depth Measurements in Digital Stereomammography Medical Physics 2002; 29(11); 2725-2734.
3) Chan HP, Goodsitt MM, Hadjiiski LM, Bailey JE, Klein K , Darner KL, Sahiner B,. Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study. Physics in Medicine and Biology 2003; 48 (22): 3721 – 3734.
4) Wei J, Chan HP, Helvie MA, Roubidoux MA, Sahiner B, Hadjiiski LM, Zhou C, Paquerault S, Chenevert T, Goodsitt MM, Correlation between Mammographic Density and Volumetric Fibroglandular Tissue Estimated on Breast MR Images, Medical Physics 2004, 31: 933-942.
5) Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Chen NG, Helvie MA, Bailey JE, Roubidoux MA, Paramagul C, Blane CE, Sahiner B, Petrick NA, An Observer Study Comparing Spot Imaging Regions Selected by Radiologists and a Computer for an Automated Stereo Spot Mammography Technique. Medical Physics 2004; 31: 1558-1567.

b) Published in non-peer-reviewed journals or in conference proceedings

1) Chan HP, Goodsitt MM, Hadjiiski L, Bailey JE, Klein K, Darner KL, Paramagul C. Digital stereomammography: observer performance study of the effects of magnification and zooming on depth perception. Proc SPIE 4682; 2002: 163-166.
2) Chan HP, Goodsitt M, Helvie MA, Hadjiiski LM, Lydick JT, Roubidoux MA, Bailey J, Nees A, Blane C, Shen J, Foster M, Sahiner B. Assessment of breast lesions on stereoscopic and monoscopic digital specimen mammograms: an ROC study. Proc SPIE 5368; 2004: 428-432.
3) Goodsitt MM, Chan, HP, Lydick JT, Kayner, D, and Helvie MA. Development of an add-on asymmetric collimator for automated stereo-spot digital mammography. Proc. Of the 7th International Workshop on Digital Mammography, University of North Carolina, Chapel Hill, NC. June 18-21, 2004. (12 page paper to be published)

c) Abstracts and presentations

1) Goodsitt MM, Chan HP, Huang H, Zhou C. Automated spot mammography for improved imaging of dense breasts. Presented at the 86th Scientific Assembly and Annual Meeting of the Radiological Society of North America in Chicago, IL, November 26-December 1, 2000 ((Radiology 2000; 217(P): 346)
2) Goodsitt MM, Chan HP, Hadjiiski LM. The effect of 2X zoom on virtual cursor depth measurements in stereomammography. Presented at the 43rd Annual Meeting of the American Association of Physicists in Medicine in Salt Lake City, Utah, July 22-26, 2001.(Medical Physics 2001;28:1247)
3) Goodsitt M, Chan HP, Gandra C, Chen N, Helvie M, Klein K, Bailey J, Paramagul C. Automated Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists. Presented at the 44th Annual Meeting of the American Association of Physicists in Medicine in Montreal, Quebec, Canada, July 14-18, 2002.(Medical Physics 2002;29:1307)
4) Chan HP, Goodsitt M, Hadjiiski L, Helvie M, Bailey J, Klein K, Roubidoux M. Development of digital stereo imaging technique for mammography. Poster presentation at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, Sept. 25-28, 2002. P28-6 in Proceedings, Volume II.
5) Goodsitt MM, Chan HP, Gandra CR, Chen NG, Helive MA. Automated stereo spot mammography for improved imaging of dense breasts. Poster presentation at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, Sept. 25-28, 2002. P28-11 in Proceedings, Volume II.
6) Goodsitt MM. Automated Stereo Spot Mammography for Improved Imaging of Dense Breasts. Presented in a Symposium entitled Digital Imaging: Diagnostic Potential and Enhancing Availability at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, September 25-28, 2002. Page 30 of Program.

7) Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Helvie MA, Bailey JE, Roubidoux M, Paramagul C, Sahiner B, Petrick N. Automated Stereo Spot Mammography: A Comparison of Spot Imaging Regions Selected by Radiologists and a Computer. Presented at the 45th Annual Meeting of the American Association of Physicist in Medicine (AAPM) in San Diego, CA August 10-14, 2003. (Medical Physics 2003;30:1456)
8) Chan HP, Goodsitt MM, Hadjiiski L, Roubidoux MA, Bailey JE, Helvie MA, Lydick JT, Sahiner B. ROC study comparing radiologists' performances in evaluating breast lesions on stereoscopic and single-projection digital specimen mammograms. Presented at the 45th Annual Meeting of the American Association of Physicists in Medicine (AAPM) in San Diego, CA August 10-14, 2003. (Medical Physics 2003;30:1456)
9) Goodsitt MM, Chan, HP, Lydick JT, Kayner, D, and Helvie MA. Development of an add-on asymmetric collimator for automated stereo-spot digital mammography. Presented at the 7th International Workshop on Digital Mammography, University of North Carolina, Chapel Hill, NC. June 18-21, 2004.
10) Chan HP, Goodsitt M, Helvie MA, Hadjiiski LM, Lydick JT, Roubidoux MA, Bailey J, Nees A, Blane C, Shen J, Foster M, Sahiner B. Assessment of breast lesions on stereoscopic and monoscopic digital specimen mammograms: an ROC study. Presented at the SPIE International Symposium on Medical Imaging, San Diego, CA, February 14-19, 2004.

(8) SCIENTIFIC PERSONNEL WHO PARTICIPATED IN PROJECT

Name	Function	Degree Earned
Mitchell M. Goodsitt, Ph.D.	PI	
Heang-Ping Chan, Ph.D.	Co-investigator	
Mark A. Helvie, MD	Co-investigator and a participant in observer study	
Hao Huang, BS	Student – program to trace borders. (Transferred to Johns Hopkins University)	
Nelson G. Chen, M.S.	Student – programs to trace borders, display regions and compute overlap indices.	
Justin T. Lydick	Student – programs for display of stereo mammograms, programs to trace borders, display regions and compute overlap indices, and control stepper motors for spot collimator	Computer Science BSE expected this year
Chaitanya Rao Gandra, MS	Student- programs to trace borders, display regions and compute overlap indices	MSEE, April, 2003
John Seamans, MS	Student – tomosynthesis (Not paid from grant.)	MS Biomed Engr, April, 2004
Dennis Kayner, BS	Machinist – built spot collimator	
Janet E. Bailey, MD	Radiologist who participated in observer study (Not paid from grant.)	
Marilyn A. Roubidoux, MD	Radiologist who participated in observer study (Not paid from grant.)	
Chintana Paramagul, MD	Radiologist who participated in observer study (Not paid from grant.)	
Caroline Blane, MD	Radiologist who participated in observer study (Not paid from grant.)	
Chuan Zhou, Ph.D.	Research scientist – assisted with initial program to display digitized mammograms and trace borders of suspicious dense regions for spot imaging. (Not paid from grant.)	

Berkman Sahiner, Ph.D.	Research scientist – assisted with applying CAD program to digitized mammograms and with display of true regions of interest (Not paid from grant.)	
Nicholas A. Petrick, Ph.D.	Research scientist – assisted with applying CAD program to digitized mammograms (Not paid from grant.)	
Katherine Klein, MD	Radiologist who participated in an early observer study (Not paid from grant.)	

(9) REPORT OF INVENTIONS: NONE

(10) CONCLUSIONS

We have developed an automated stereo spot imaging technique for improved visualization of dense regions in mammograms. This was an evolution from our original goal, which was to develop an automated spot compression and collimation technique. The technique was modified based on initial investigations with compressible breast simulating phantoms. These studies showed spot compression can squeeze masses and overlying or underling dense tissue closer together rather than produce the desired separation; whereas 3D imaging techniques such as stereo mammography or tomosynthesis should yield improved conspicuity of superimposed objects. In addition the 3D techniques should be more practical since they do not require automated spot compression and the use of a breast restraining device during the changeover to the spot compression paddle. Observer studies were performed comparing the suspicious regions that different radiologists would choose for spot imaging with those selected by a CAD computer program. We found there was significant variation in the overlap between the radiologist and CAD selected ROIs due mostly to variations in the numbers of radiologist selected ROIs per image and the sizes of the radiologist selected ROIs, with greater numbers and larger sizes having greater probabilities of overlap. We also performed a study comparing the radiologist and CAD selected regions to true mass regions determined by an experienced MQSA certified radiologist using mammograms, biopsy images and pathology results. The CAD-selected ROIs had a 73% overlap with the TROIs, which was comparable with the average overlap result of 76.8% +/- 10% for the 5 experienced radiologists who participated in our observer study. Thus, CAD determined ROIs could potentially be useful for a screening technique that includes stereo spot mammography. We determined the geometric requirements for the automated spot collimator to insure the suspicious region in a full-field digital mammogram is adequately covered during stereo spot image acquisition. These requirements were implemented in software program that controlled the positions of the spot collimator blades. Tests on phantoms containing simulated masses showed the software worked properly.

A modular breast phantom was designed containing slabs with simulated masses, dense regions and microcalcifications that could be arranged in various ways to produce conventional mammograms with overlapping masses and densities . Although the manufacturer of this phantom was not very successful at meeting our design requirements, we did find it was difficult to distinguish subtle overlapping masses even with spot stereo mammography of this phantom. We performed a preliminary study of another 3D imaging technique tomosynthesis and found it to be very promising. We plan to continue our spot mammography work on a dedicated tomosynthesis device that is being developed for one of our other research projects. We should be able to compare spot and full-field stereo mammography with spot and full-field tomosynthesis with this device to determine which is best at discriminating overlapping tissues for breast cancer detection..

So What

A key limitation of conventional x-ray mammography is the inability to optimally image regions of dense and overlapping tissue. The new full-field digital mammography systems reduce but do not entirely eliminate this problem. Additional views including spot imaging are still employed with the digital systems to better analyze suspicious tissue regions. The system we are developing adds 3 features that will further optimize spot imaging of suspicious tissue. First, computer vision techniques will be employed to automatically find suspicious regions that warrant spot imaging. Second, the system will automatically collimate the x-ray beam to the suspicious region. This will reduce the amount of x-ray scatter that strikes the detector thereby reducing image noise. It should also restrict the region in which the automated optimization of parameters (AOP (kVp, target, and filter)) and phototiming are determined, which may improve the penetration of the spot region. Third, the spot image will be acquired using a stereoscopic method. This will produce a 3D image allowing the radiologist to see the depth separation between lesions and overlapping tissues and allowing the radiologist to see through any "cloud" of dense tissue. Stereoscopic imaging also permits improved visualization of the locations and 3D shapes of lesions and microcalcification clusters, and of lesion borders, all of which are important to the radiologists in making their diagnoses. Finally, the automated spot collimation stereo mammography technique is performed with the breast in the same location as in the full-field mammogram, eliminating the guesswork associated with breast re-positioning in conventional spot compression mammography. Thus, in theory, the automated spot collimation, stereomammography technique should be of great benefit. The challenge will be to develop a convenient and practical system that could be employed in screening, which is the long-term goal of our project

(11) REFERENCES

1) Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Chen NG, Helvie MA, Bailey JE, Roubidoux MA, Paramagul C, Blane CE, Sahiner B, Petrick NA, An Observer Study Comparing Spot Imaging Regions Selected by Radiologists and a Computer for an Automated Stereo Spot Mammography Technique. Medical Physics 2004; 31: 1558-1567

2) Petrick N, Sahiner B, Chan HP, Helvie MA, Paquerault S, and. Hadjiiski LM, Breast cancer detection: Evaluation of a mass detection algorithm for computer-aided diagnosis: Experience in 263 patients, Radiology 2002; 224: 217-224

3) Chan HP, Goodsitt MM, Hadjiiski L, Bailey JE, Klein K, Darner KL, Paramagul C. Digital stereomammography: observer performance study of the effects of magnification and zooming on depth perception. Proc SPIE 4682. 2002: 163-166.

4) Goodsitt MM, Chan HP, Gandra CR, Chen NG, Helive MA. Automated stereo spot mammography for improved imaging of dense breasts. Poster presentation at the Era of Hope, Department of Defense Breast Cancer Research Program Meeting in Orlando, Florida, Sept. 25-28, 2002. P28-11 in Proceedings, Volume II.

5) Niklason LT, Christian BT, Niklason LE, Kopans DB, Castleberry DE, Opsahl-Ong BH , Landberg CE, Slanetz PJ, Giardino AA, Moore R, Albagl Di D, DeJuleMC, Fitzgerald PF, , et al,. "Digital tomosynthesis in breast imaging," Radiology 1997; 205: 399-406.

6) Goodsitt MM, Chan HP, Darner K and Hadjiiski LM. The Effects of Stereo Shift Angle, Geometric Magnification, and Display Zoom on Depth Measurements in Digital Stereomammography Medical Physics 2002; 29(11); 2725-273.

7) Chan HP, Goodsitt MM, Hadjiiski LM, Bailey JE, Klein K , Darner KL, Sahiner B,. Effects of Magnification and Zooming on Depth Perception in Digital mammography: An Observer Performance Study. Physics in Medicine and Biology 2003; 48 (22): 3721 – 3734.

8) Wu T, Stewart A, Stanton M, McCauley T, Philips W, Kopans DB et al. Tomographic mammography using a limited number of low-dose cone-beam projections. Medical Physics 2003; 30(3): 365-380.

9) Dobbins JT and Godfrey DJ. Digital x-ray tomosynthesis: current state of the art and clinical potential. Phys. Med. Biol. 2003; 48: R65-R106.

(12) APPENDIX

The following publications as a result of this grant in the final year are enclosed with this report.

1 Wei J, Chan HP, Helvie MA, Roubidoux MA, Sahiner B, Hadjiiski LM, Zhou C, Paquerault S, Chenevert T, Goodsitt MM, Correlation between Mammographic Density and Volumetric Fibroglandular Tissue Estimated on Breast MR Images, Medical Physics 2004, 31: 933-942.

2 Goodsitt MM, Chan HP, Lydick JT, Gandra CR, Chen NG, Helvie MA, Bailey JE, Roubidoux MA, Paramagul C, Blane CE, Sahiner B, Petrick NA, An Observer Study Comparing Spot Imaging Regions Selected by Radiologists and a Computer for an Automated Stereo Spot Mammography Technique. Medical Physics 2004; 31: 1558-1567

3 Goodsitt MM, Chan, HP, Lydick JT, Kayner, D, and Helvie MA. Development of an add-on asymmetric collimator for automated stereo-spot digital mammography. Proc. of the 7th International Workshop on Digital Mammography, University of North Carolina, Chapel Hill, NC. June 18-21, 2004. (12 page paper to be published)

4 Chan HP, Goodsitt M, Helvie MA, Hadjiiski LM, Lydick JT, Roubidoux MA, Bailey J, Nees A, Blane C, Shen J, Foster M, Sahiner B. Assessment of breast lesions on stereoscopic and monoscopic digital specimen mammograms: an ROC study. Proc SPIE 5368; 2004: 428-432.

Correlation between mammographic density and volumetric fibroglandular tissue estimated on breast MR images

Jun Wei,[a] Heang-Ping Chan, Mark A. Helvie, Marilyn A. Roubidoux, Berkman Sahiner, Lubomir M. Hadjiiski, Chuan Zhou, Sophie Paquerault, Thomas Chenevert, and Mitchell M. Goodsitt
Department of Radiology, University of Michigan, Ann Arbor, Ann Arbor, Michigan 49109

(Received 18 June 2003; revised 26 November 2003; accepted for publication 21 January 2004; published 26 March 2004)

Previous studies have found that mammographic breast density is highly correlated with breast cancer risk. Therefore, mammographic breast density may be considered as an important risk factor in studies of breast cancer treatments. In this paper, we evaluated the accuracy of using mammograms for estimating breast density by analyzing the correlation between the percent mammographic dense area and the percent glandular tissue volume as estimated from MR images. A dataset of 67 cases having MR images (coronal 3-D SPGR T1-weighted pre-contrast) and corresponding 4-view mammograms was used in this study. Mammographic breast density was estimated by an experienced radiologist and an automated image analysis tool, Mammography Density ESTimator (MDEST) developed previously in our laboratory. For the estimation of the percent volume of fibroglandular tissue in breast MR images, a semiautomatic method was developed to segment the fibroglandular tissue from each slice. The tissue volume was calculated by integration over all slices containing the breast. Interobserver variation was measured for 3 different readers. It was found that the correlation between every two of the three readers for segmentation of MR volumetric fibroglandular tissue was 0.99. The correlations between the percent volumetric fibroglandular tissue on MR images and the percent dense area of the CC and MLO views segmented by an experienced radiologist were both 0.91. The correlation between the percent volumetric fibroglandular tissue on MR images and the percent dense area of the CC and MLO views segmented by MDEST was 0.91 and 0.89, respectively. The root-mean-square (rms) residual ranged from 5.4% to 6.3%. The mean bias ranged from 3% to 6%. The high correlation indicates that changes in mammographic density may be a useful indicator of changes in fibroglandular tissue volume in the breast. © *2004 American Association of Physicists in Medicine.* [DOI: 10.1118/1.1668512]

Key words: mammography, breast density, MR images, correlation

I. INTRODUCTION

Studies have shown that there is a strong positive correlation between breast parenchymal density imaged on mammograms and breast cancer risk.[1–3] The relative risk is estimated to be about 4 to 6 for women whose mammograms have parenchymal densities over 60% of the breast area, as compared to women with less than 5% densities. Other cohort studies[4–13] also found that breast cancer risk in the category with the most extensive dense tissue was 1.8 to 6 times as high as that in the category with the least extensive dense tissue. Mammographic density as the risk indicator is greater than almost all other risk factors of breast cancer.[2,14] Although there is no direct evidence that changes in mammographic breast densities will result in changes in breast cancer risk, the strong correlation between breast density and breast cancer risk has prompted researchers to use mammographic density for monitoring the effects of intervention as well as for studying breast cancer etiology.[14–17]

A number of researchers have investigated image analysis techniques to estimate breast density.[15,18–28] The common approaches are to analyze the textural pattern or the percentage of mammographic densities relative to the breast area. It has been found that the texture measures were correlated with parenchymal density patterns but they appeared to be less sensitive measures of relative risk than the percent dense area.[1,25,29] In current practice, breast density is estimated mainly by radiologists' visual judgment of the fibroglandular tissue imaged on mammograms following the Breast Imaging—Reporting and Data System (BI-RADS) lexicon.[30,31] Because of the qualitative and subjective nature of visual judgment, there are large intraobserver and interobserver variations in the estimated breast density. The large variability may reduce the observed correlation between breast cancer risk and breast density. It may also reduce the sensitivity of studies using mammographic density for monitoring the effect of risk modifying treatments. We have developed an automated image analysis system, Mammographic Density ESTimator (MDEST), to assist radiologists in estimating breast density on mammograms. A computerized analysis is expected to increase the reproducibility and consistency in the estimation of mammographic density, thereby improving the accuracy of the related studies. In our previous study, we have found that the percent mammographic density segmented by MDEST agreed closely with that estimated by radiologists' interactive thresholding.[32]

The high correlation between breast cancer risk and breast

 0094-2405/2004/31(4)/933/10/$22.00

density indicates that breast cancer risk may be closely related to the volume of glandular tissue in the breast. Among the modalities available for breast imaging at present, magnetic resonance (MR) imaging is likely to be the most accurate method for volumetric dense tissue estimation because fibroglandular tissue and adipose tissue can be well distinguished in MR images when a proper image acquisition technique is used.[33] However, MR imaging is expensive, making it difficult to use MR imaging as a routine monitoring tool.[33,34] On the other hand, a mammogram is a two-dimensional (2-D) projection image of a three-dimensional (3-D) object. The area of dense tissue measured on a mammogram is not an accurate measure of the volume of fibroglandular tissue in the breast because no thickness information is used. However, mammography is a widely available low cost procedure that may be used for monitoring breast density change during preventive and interventional treatment or other studies. Women who participate in screening will also have mammograms readily available for retrospective review. Therefore, mammography will most likely be the method of choice for breast density estimation.

In this study, we investigated the correlation between the volumetric fibroglandular tissue in the breast and the projected breast dense area on mammograms by analyzing the percent volumetric fibroglandular tissue in MR breast images and the percent dense area in corresponding mammograms. Our purpose in this study is not to evaluate the usefulness of either MR fibroglandular tissue volume or mammographic density as an indicator for breast cancer risk, which have been studied by other investigators. Rather, we used the MR breast images to estimate the volumetric fibroglandular tissue in the breast and explored the reason that a change in mammographic density (2-D) can be used as an indicator of breast density change (3-D). These comparisons will provide a better understanding of their relationship, and may lead to improved methods for utilizing mammographic density as a surrogate marker for breast cancer risk.

II. MATERIALS AND METHOD

A. Dataset

In a previous study, gadolinium contrast enhanced MR dynamic imaging was employed to characterize malignant and benign breast lesions. A dataset was collected with IRB approval which included MR images and corresponding mammograms acquired between detection and before biopsy for a given patient. In the MR study, several series of images were acquired for each patient. Patients were scanned prone using a commercial dual phased-array breast coil. The imaging protocol included a series was the coronal 3-D T1-weighted pre-contrast series (coronal sections 2–5 mm thick, 32 slices; 3-D Spoiled Gradient-Recalled Echo (SPGR); TE = 3.3 ms; TR = 10 ms, Flip = 40°, matrix = 256×128, FOV = 28–32 cm right/left, 14–16 cm superior/inferior, scan time = 2 min 38 sec). This 3-D SPGR sequence produces full volume coverage of both breasts with contiguous image sections. The dense parenchyma and fat tissue are well separated with this heavily T1-weighted acquisition. We used a set of 67 patients to study the correlation between the 2-D projected percentage of dense area on a mammogram and the percentage of dense tissue volume estimated from the 3-D MR images.

The mammograms consisting of the craniocaudal (CC) view and the mediolateral oblique (MLO) view of both breasts of the patient were digitized with a LUMISYS 85 laser film scanner at a pixel size of 50 μm×50 μm. The digitizer has a gray level resolution of 12 bits and a nominal optical density (O.D.) range of 0 to 4. For density segmentation, it is not necessary to use very high-resolution images. To reduce processing time, the full resolution mammograms were first smoothed with a 16×16 box filter and subsampled by a factor of 16, resulting in 800 μm×800 μm images for this study.

B. Estimation of fibroglandular tissue volume on MR images

Since it is not our intention to routinely segment MR images for breast density estimation, we did not attempt to develop an automated method for this application. Our algorithm for segmentation of volumetric fibroglandular tissue on MR images used a semi-automatic method. The computer performed an initial segmentation. A graphical user interface (GUI) was developed to allow a user to review the segmentation of every slice and make modifications if necessary. The method consists of four steps. First, the breast boundary was detected automatically on each slice. A deformable model and manual modification were used to correct for incorrectly detected boundaries that usually occurred in slices near the chest wall where there were no well-defined breast boundaries. Because of inhomogeneity of the breast coil sensitivity, the signal intensity in the breast region was not uniform across the field of view. A background correction technique that estimated the low frequency background from the gray levels along the breast boundary was developed to reduce this systematic nonuniformity. Manual interactive thresholding of the gray level histogram in the breast region was then used to separate the fibroglandular from the fatty region. Morphological erosion was used to exclude the skin voxels along the breast boundary. Finally, the volume of fibroglandular tissue was calculated by integration over all slices containing the breast. A flow chart of our algorithm is shown in Fig. 1.

C. Breast boundary detection

A two-step algorithm was developed for the detection of breast boundary on each slice. First, we used a seeded pixel thresholding algorithm (SPTA) for the initial assessment of a breast boundary. Second, a 2-D active contour algorithm further refined the boundary. For slices close to the chest wall where no clear boundary can be seen, manual modification was used to outline an estimated boundary.

The SPTA determined the optimal threshold by iteratively partitioning the MR image into two parts and using the gradient value along the boundary of the partition as a guide in optimizing the threshold. First, the center of gravity was se-

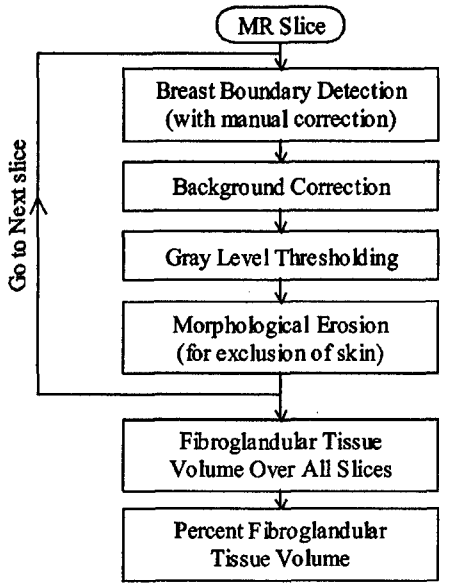


FIG. 1. The flow-chart for the segmentation of the fibroglandular tissue on MR images.

lected as the starting pixel on each slice. The gray level of the starting pixel was used as a threshold to create a binary partition of the image in which all pixels greater than the threshold were set to one and all other pixels were set to zero. Second, the gradient value of each pixel on the boundary of the binary partition was calculated by applying the Sobel filter to the original image. The gradient assessment for this particular binary partition was defined as the average gradient magnitude of these boundary pixels. The threshold value was reduced to zero in a stepwise manner. The partition for each threshold value was created and the gradient assessment for each partition was calculated as described above. The partition with the maximum gradient assessment was considered to be the initial segmentation result for the breast, and the boundary of this partition was considered to be the initial breast boundary.

After the initial segmentation, a deformable contour method was used to further refine the boundary. The movement of the boundary pixel was controlled by an energy function which consisted of internal energy and external energy. The internal energy components used in this study were the continuity and curvature of the contour, as well as the homogeneity of the segmented partition. The external energy components were the negative of the smoothed image gradient magnitude, and a balloon force that exerted pressure at a normal direction to the contour. The energy function was defined as the following:

$$E=\sum_{c=1}^{N}[E_{\text{inter}}(c)+E_{\text{exert}}(c)], \tag{1}$$

where E_{inter} and E_{exert} are the internal energy and the external energy, respectively, as defined in Eq. (2) and Eq. (3):

$$E_{\text{inter}}=w_{\text{curv}}E_{\text{curv}}(c)+w_{\text{cont}}E_{\text{cont}}(c)+w_{\text{hom}}E_{\text{hom}}, \tag{2}$$

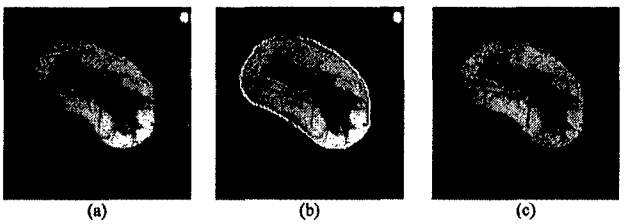


FIG. 2. An example of the first three processing blocks in Fig. 1. (a) Original MR slice; (b) automatically-detected breast boundary superimposed on the image; and (c) the background-corrected image.

$$E_{\text{exert}}=w_{\text{grad}}E_{\text{grad}}(c)+w_{\text{bal}}E_{\text{bal}}(c), \tag{3}$$

where curv, cont, grad, bal, hom denoted curvature, continuity, gradient, balloon force and homogeneity, respectively, and each energy term was associated with a weight, w. The detailed definition for each term can be found in the literature.[35] An example of a MR slice of a breast is shown in Fig. 2(a), and the segmented boundary is shown in Fig. 2(b). Note that the two breasts of a patient were scanned together but each breast was analyzed separately.

D. Background correction

To reduce the nonuniformity of the MR signal intensity in the breast region, a background correction technique[36] using the pixel values around the segmented breast region was employed. For a given pixel (i,j) inside the breast region, the gray value of the background image was estimated as shown in Eq. (4):

$$B(i,j)=\left[\frac{L}{d_l}+\frac{R}{d_r}+\frac{U}{d_u}+\frac{D}{d_d}\right]\Big/\left[\frac{1}{d_l}+\frac{1}{d_r}+\frac{1}{d_u}+\frac{1}{d_d}\right], \tag{4}$$

where L, R, U and D are the average gray values inside a breast background estimation region (BBER) centered at the left, right, upper and lower pixels on the breast boundary, respectively. A BBER was defined as the intersection of a 21×21-pixel box and the breast region. The center pixels for the left and right boxes were the intersection points between the breast boundary and a horizontal line passing through the given pixel (i,j). Similarly, the upper and lower center pixels for the upper and lower boxes were the intersection points between the breast boundary and a vertical line passing through the given pixel (i,j). Only the pixels that were within the intersected area between the 21×21-pixel box and the breast region were included in the definition of the BBER and the calculation of the average gray value. The contributions of the average gray levels to the background pixel (i,j) were inversely weighted by their distances d_l,d_r,d_u,d_d from the given pixel (i,j). An example of the background corrected image is shown in Fig. 2(c).

E. Segmentation of fibroglandular tissue

We developed a GUI that allowed the user to perform a combination of manual and automatic operations to segment the breast boundary and the fibroglandular tissue on the MR

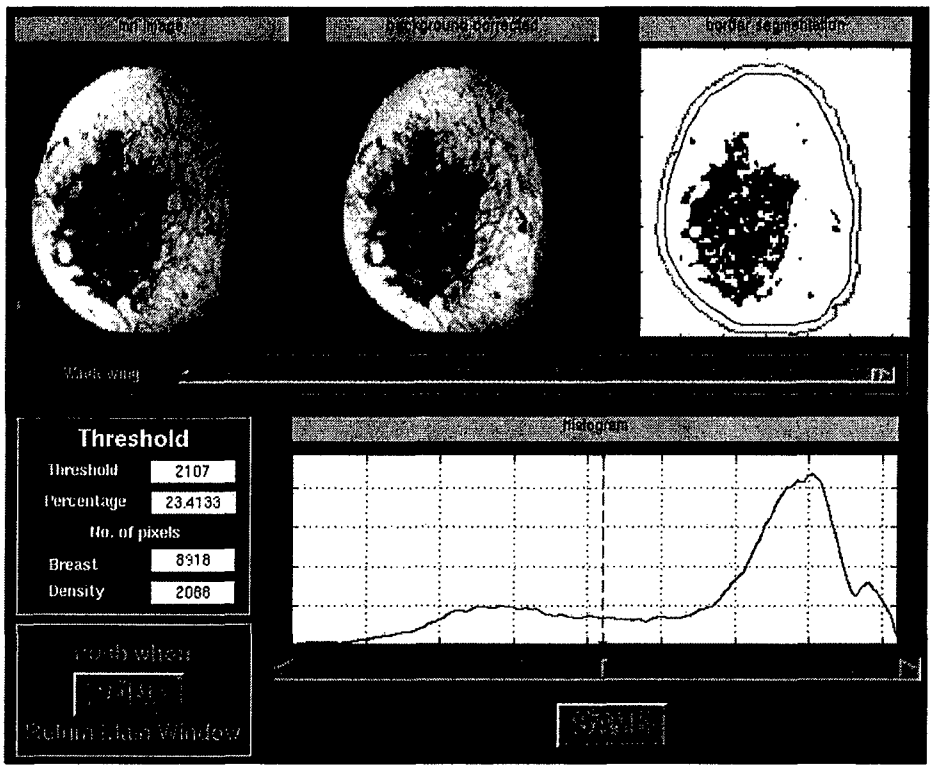


FIG. 3. The graphic user interface for the segmentation of the fibroglandular tissues on the MR slice. The upper row shows the original MR slice (left), the background-corrected image (middle) and the segmented binary image (right). The segmented image responds to the reader's adjustment of the gray level threshold (lower row) in real time so that the reader can choose the appropriate threshold by inspecting the segmented image visually. The dark area in the segmented image indicates the fibroglandular tissue and the white area indicates the adipose tissue. The inner line along the breast boundary is the boundary obtained by morphological erosion to exclude the skin voxels for calculating the fibroglandular tissue volume.

images. The first window (not shown) displayed the MR series and the corresponding mammogram of each breast to give the user an overview of the breast. The segmentation of the fibroglandular tissue on each MR slice was processed in the second window, shown in Fig. 3. The original MR slice, the corresponding background corrected image and the segmented binary image were shown in the upper part of the window. At the lower part of the window, the histogram of the voxel values in the breast region was shown. The user performed interactive thresholding on the histogram and the segmented binary image corresponding to the chosen threshold was displayed in real time in the upper part. If the breast boundary, which was automatically segmented by the computer initially, had to be corrected, the user could go to the third window and manually move the apices of the polygon outlining the boundary. The voxels contributed by the nipple were excluded. On the slices containing breast skin that had voxel values similar to those of fibroglandular tissue, a morphological erosion operation was applied to the breast boundary to exclude the skin voxels from the calculation of the fibroglandular tissue volume in the slice. The size of the structuring element could be selected interactively on the fourth window and the eroded boundary was displayed instantly for a chosen erosion operation. The user might again change the structuring element if the erosion result of the previous choice was deemed unsatisfactory. Since the eroded boundary only marked the region within which the fibroglandular voxels would be summed and would not be used for the calculation of the breast volume, as described below, it did not need to be precise as long as it excluded the skin voxels while not excluding the fibroglandular voxels.

F. MR fibroglandular tissue volume

After the fibroglandular tissue was segmented for each slice, the total number of voxels containing the fibroglandular tissue was obtained as a summation of these voxels over all slices of the breast. The total volume of the breast was obtained as the summation of the voxels enclosed by the breast boundary before morphological erosion. The ratio of these two volumes provided the percent volumetric fibroglandular tissue in the breast.

G. Mammographic density segmentation

We have previously developed an automated method for segmentation of the dense fibroglandular area on mammograms. The method, referred to as the Mammographic Density ESTimator (MDEST) was described in detail elsewhere.[32] In brief, the breast boundary on the digitized mammogram is tracked. A dynamic-range compression technique reduces the gray level range of the breast area. By analyzing the shape of the gray level histogram, a rule-based classifier classifies the breast density into one of four classes. Typically, a Class I breast is almost entirely fat; it has a single narrow peak on the histogram. A Class II breast contains scattered fibroglandular densities. Its histogram has two main peaks, with the smaller peak on the right of the bigger one. A Class III breast is heterogeneously dense. Its histogram also has two peaks, but the smaller peak is on the left of the bigger one. A Class IV breast is extremely dense. Its histogram has mainly a single dominant peak, but the peak is wider compared with the peak in the Class I histogram. A second smaller peak sometimes occurs on the left of the

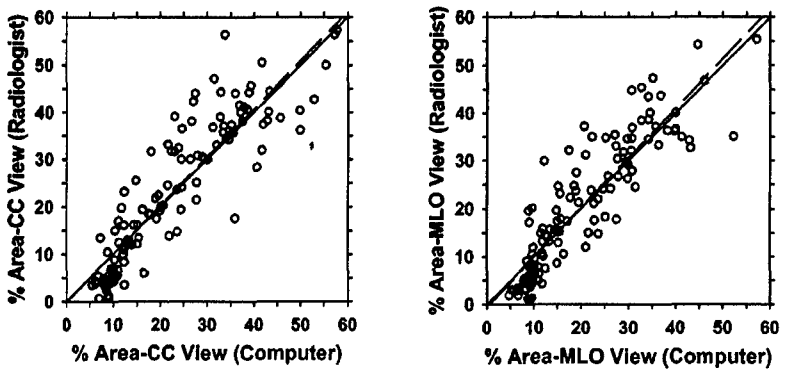


FIG. 4. A comparison of the percent mammographic density obtained from interactive thresholding by an MQSA-qualified radiologist and that estimated by our automated MDEST computer program. (a) CC view, correlation coefficient=0.90, rms residual=6.7, mean difference=0.3; (b) MLO view, correlation coefficient =0.89, rms residual=6.1, mean difference=0.4. Dashed line: linear regression of the data; solid line: diagonal.

main peak. Based on the histogram shape, a threshold is automatically calculated to separate the dense and fatty pixels. The mammographic density was estimated as the percentage of fibroglandular tissue area relative to the total breast area. For MLO view mammograms, the pectoral muscle is detected and excluded from the density area or breast area calculations. In our previous work, the performance of MDEST was verified by comparison with manual segmentation by 5 breast imaging radiologists using a dataset of 260 mammograms from 65 patients that were different from the cases used in the current study. We found that the correlation between the computer-estimated percent dense area and the average segmentation by the 5 radiologists was 0.94 and 0.91, respectively, for CC and MLO views, with a mean bias of less than 2%.

MDEST was applied to the mammograms of the 67 patients used in this study. The percent dense area on mammograms was estimated for the CC-view and the MLO-view mammogram of each breast separately. In addition, an MQSA-qualified radiologist also segmented the dense area by interactive thresholding for each mammogram. The correlation between the mammographic density obtained by manual and automatic segmentation is shown in Figs. 4(a) and 4(b) for the CC view and MLO view, respectively. The correlation coefficients for the CC view and MLO view were 0.90 and 0.89, respectively. The mammographic densities estimated by automatic and manual segmentation were compared with the percent volumetric fibroglandular tissue on MR images as described below.

H. Observer experiments

We performed an experiment to evaluate the variability of the estimated % volumetric fibroglandular tissue due to the uncertainty in the determination of the starting slice of the breast at the chest wall. The starting slice affected the estimation of the breast volume that was calculated by integrating from the starting slice to the anterior of the breast. Twenty-three MR cases from the dataset were randomly selected for this observer experiment. There were a total of 41 breasts because some cases had only one breast. For this subset of cases, each radiologist was asked to select the starting slice from the MR images for each breast. The estimated % volumetric fibroglandular tissue calculated with all available slices was then compared to that calculated with the selected starting slice.

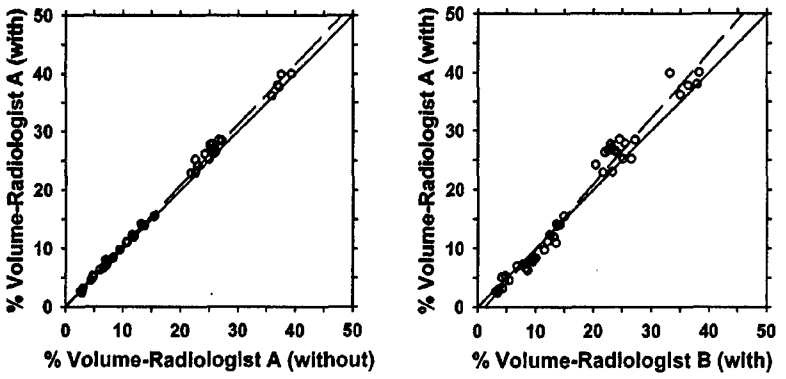
We also performed observer experiments to evaluate the inter-observer variations in the segmentation of fibroglandular tissue using the semi-automatic method. Two MQSA-qualified radiologists performed the segmentation of the fibroglandular tissue on the MR images of the 41 breasts using the semi-automatic method implemented with the GUI. A Ph.D. researcher who was trained by these radiologists also performed the segmentation independently with the GUI.

After verifying the consistency of segmentation by these observers, the trained Ph.D. completed the segmentation of all MR cases. The correlation between percent volumetric fibroglandular tissue on MR images and percent dense area on mammograms was then examined for the entire dataset.

III. RESULTS

A. Effect of selection of the starting slice

Figure 5(a) shows the correlation of the % volumetric fibroglandular tissue calculated using all available slices for the breast with that calculated using the selected starting slice by radiologist A for the 41 breasts. The correlation coefficient was 0.999. To compare the difference between their results, the mean difference and the root-mean-square (rms) residual, which is the residual from the linear least-squares-fitted line, were also calculated. The mean difference was 0.7 and the rms residual was 0.6. The result is similar for radiologist B (not shown), with a correlation coefficient of 0.999, a mean difference of 0.4 and a rms residual of 0.4. The correlation between the % volumetric fibroglandular tissue calculated using the selected starting slice by radiologist A with that calculated using the selected starting slice by radiologist B was also very high with a correlation coefficient of 0.988, a mean difference of 0.7 and a rms residual of 1.8, as shown in Fig. 5(b). These comparisons indicated that the variability in the selection of the starting slice of the breasts did not have a strong influence on the % volumetric fibroglandular tissue. We therefore used all available slices in the MR dataset for each breast in the following analyses.



(a) (b)

FIG. 5. (a) A comparison of the percent fibroglandular tissue volume calculated using the selected starting slice with that calculated using all available slices for radiologist A, correlation coefficient=0.999. (b) A comparison of the percent fibroglandular tissue volume calculated using the selected starting slice by radiologist B with that by radiologist A, correlation coefficient =0.988, Dashed line: linear regression of the data; solid line: diagonal.

B. Inter-observer variation between radiologists

Figure 6(a) shows the comparison of the percent volumetric fibroglandular tissues on MR images segmented by two radiologists for the 41 breasts. The correlation between the segmentation results of the two radiologists is 0.99. The mean difference was found to be 0.3 and the rms residual was 1.6.

C. Inter-observer variation between radiologists and trained Ph.D.

Figure 6(b) shows the comparison of the percent volumetric fibroglandular tissues segmented by the trained Ph.D. against that segmented by radiologist A. A similar result was obtained by comparing the percent volumetric tissue segmented by the trained Ph.D. and that segmented by radiologist A except that the data points were even closer to the diagonal (not shown). The correlation between the result of the trained Ph.D. and the results of both radiologists was 0.99. The corresponding mean differences were -0.8 and -0.4, respectively, and the rms residuals were 1.4 and 1.5, respectively.

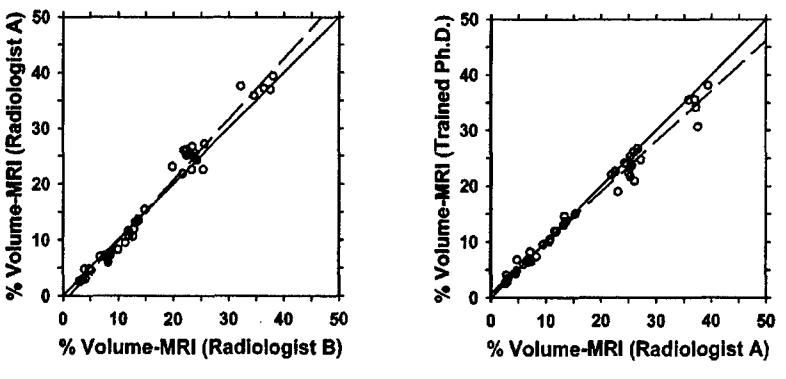
D. Correlation between percent volumetric fibroglandular tissue on MR images and percent mammographic density

The percent volumetric fibroglandular tissue on MR images was compared with the percent dense area on CC- and MLO-view mammograms. After verifying that the difference in segmentation between the trained Ph.D. and the radiologists was similar to the interobserver variations between the two experienced radiologists, the trained Ph.D. completed the segmentation of the entire dataset.

Figure 7 shows the comparison of the percent volumetric fibroglandular tissue on MRI and the percent mammographic density segmented by a radiologist. The percent areas on CC- and MLO-view mammograms are higher than the percent volume on MR images with a mean difference of 5.7% and 3.0%, respectively.

Figure 8 shows the comparison of the percent volumetric fibroglandular tissue on MRI and the percent mammographic density segmented by MDEST. The percent areas on CC- and MLO-view mammograms segmented by the computer are higher than the percent volume on MR images with a mean difference of 5.3% and 2.6%, respectively.

The correlation coefficients, the mean differences and the rms residuals between the percent volumetric fibroglandular tissue on MR images and percent dense area on mammograms are compared in Table. I. The correlation between the percent volume on MR images and percent area on mammograms of the fibroglandular breast tissue is high, ranging from 0.89 to 0.91. Although it is not expected that the values of percent volume agree with the values of percent area, their mean differences range only from 3% to 6% and the rms residual range from 5.4 to 6.3.



(a) (b)

FIG. 6. A comparison of the segmentation of fibroglandular tissue from MR images between two observers: (a) two experienced MQSA-qualified radiologists, correlation coefficient=0.99. (b) The trained Ph.D. and Radiologist A, correlation coefficient=0.99. The correlation between the trained Ph.D. and Radiologist B is also 0.99 but the data points were very close to the diagonal and is not shown. The % volumetric fibroglandular tissue was calculated using all available slices. Dashed line: linear regression of the data; solid line: diagonal.

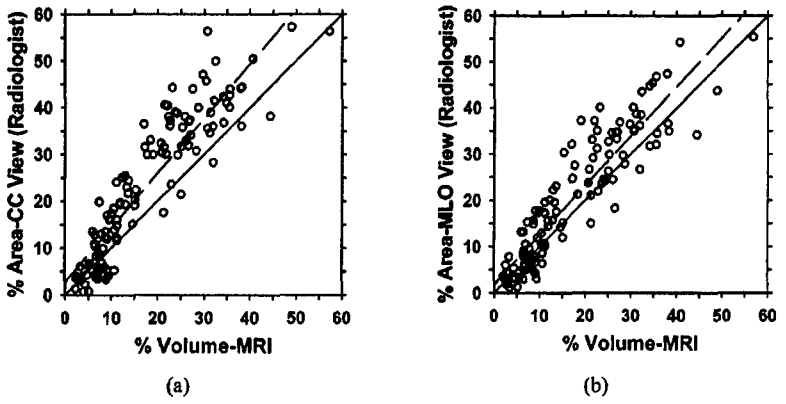


FIG. 7. A comparison of the percent fibroglandular tissue volume on MR images and the percent dense area on mammograms segmented by an experienced radiologist. (a) CC view, correlation coefficient=0.91; (b) MLO view, correlation coefficient=0.91. Dashed line: linear regression of the data; solid line: diagonal.

IV. DISCUSSION

Our purpose in this paper was to investigate the relationship between the percent dense area on mammogram and the percent fibroglandular tissue volume on MR image. We found a direct correlation between mammographic density and MR volumetric density (Fig. 7 and Fig. 8). The correlation coefficients between the percent area on a mammogram and the percent volume on MR images are high at 0.89 and 0.91. These results are more promising than those found in previous studies that attempted to correlate percent dense area on mammograms with MR information. Graham *et al.*[33] investigated the relationship between percent density (projected dense area) on mammogram and two objective MR parameters of breast tissue, relative water content and mean T2 relaxation. Their results with 45 cases showed a positive correlation between percent density and relative water content (Pearson correlation coefficient=0.79) and a negative correlation between percent density and mean T2 value (Pearson correlation coefficient=−0.61). Another study by Lee *et al.*[34] analyzed fatty and fibroglandular tissue in different age groups to compare x-ray mammography with T1-weighted MR images. Their study with 40 cases indicated that the correlation between the two techniques is 0.63 when the fat content was more than 45%. However, the correlation coefficient decreased to 0.34 when their analysis included only dense breasts.

It may be noted that although MR imaging is currently the most accurate method for estimating the volumetric fibroglandular tissue in the breast, it is still not the ideal tool. Fibrous tissue and glandular tissue are not well separated with current MR imaging techniques. Since the amount of glandular tissue in the breast is the important factor relating to breast cancer risk, further studies are warranted for differentiating the glandular and the fibrous components of the imaged volume. The correlation between the percent glandular tissue volume and percent projected dense area on a mammogram will be a more reliable indicator of the usefulness of mammographic density analysis.

The density on mammograms is a 2-D projected area of the fibroglandular tissues. The percent dense area is not expected to be equal in value to the percent volume. The mean differences between the percent volume and the percent area on CC- and MLO-views, as determined by the radiologist's interactive segmentation, are 5.7 and 3.0, respectively (Table I), with the percent dense area values being higher. We also investigated the rms residual between the percent volume and the percent area when the relationship between them was assumed to be linear. The rms residual between the percent volume and the percent area on CC- and MLO-views are 6.3 and 5.6, respectively (Table I), relative to the straight line obtained from linear least squares fits to the data. One possible factor that may contribute to a higher value of percent dense area on mammograms than the percent volume value on MR images is that the tissue volume imaged by the two modalities is somewhat different. The MR images include more tissue near the chest wall, which is mainly retroglan-

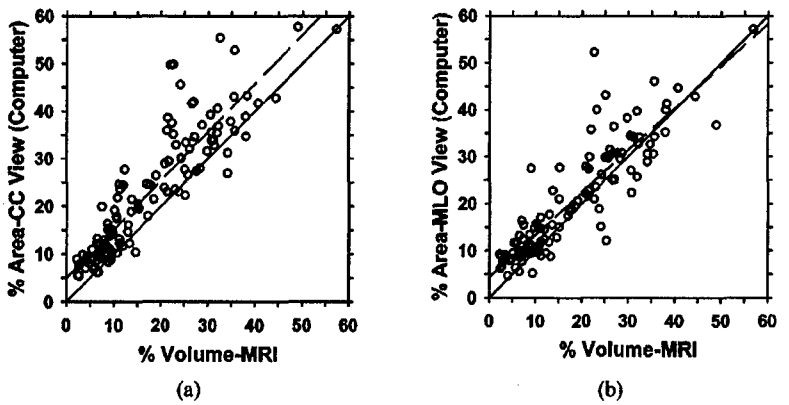


FIG. 8. A comparison of the percent volume on MR images and the percent area on mammogram segmented by our automated MDEST computer program. (a) CC view, correlation coefficient=0.91; (b) MLO view, correlation coefficient=0.89. Dashed line: linear regression of the data; solid line: diagonal.

TABLE I. Statistic analysis of the relationship between percent fibroglandular tissue volume on breast MR images and percent dense area on mammograms segmented by radiologist and MDEST.

	Radiologist		*Computer (MDEST)*	
	CC vs MRI	MLO vs MRI	CC vs MRI	MLO vs MRI
Correl. coeff.	0.91	0.91	0.91	0.89
rms residual	6.3	5.6	5.8	5.4
Mean diff.	5.7	3.0	5.3	2.6

TABLE II. An analysis of the relationship between percent fibroglandular tissue volume (x) on breast MR images and percent dense area (Y) on mammograms segmented by radiologist using three mathematical models. m, k, k_1 and k_2 are constants determined by least squares curve fitting.

Mathematical model		$Y=kx^{2/3}$	$Y=kx^m$	$Y=k_1x^m+k_2$
CC vs MRI	Least squares Fit	$Y=0.82x^{2/3}$	$Y=1.03x^{0.83}$	$Y=1.02x^{0.48}-0.19$
	rms residual	6.5	6.0	5.6
	Coefficient of determination	0.82	0.85	0.87
MLO vs MRI	Least squares Fit	$Y=0.73x^{2/3}$	$Y=0.96x^{0.86}$	$Y=0.90x^{0.60}-0.09$
	rms residual	6.0	5.5	5.3
	Coefficient of determination	0.80	0.84	0.85

dular adipose tissue, than a mammogram does, thus reducing the percentage of fibroglandular tissue volume. The reduction in the percent volume values, however, is relatively small, as found in our study evaluating the effects of selecting starting slices for volume calculation (Fig. 5). The main difference may therefore be attributed to the geometric relationship between the volume and the projected 2-D area, explained later.

Geometrically, we do not expect the relationship between volume and its projected 2-D area to be linear. In a hypothetical situation such that the dense tissue volume is a sphere (volume$=4/3\ \pi r^3$) enclosed inside a concentric spherical shell of fatty tissue volume, the percent projected 2-D area (area$=\pi r^2$) of the inner sphere relative to the outer sphere is equal to the percent volume to the power of 2/3. The relationship between the percent area and the percent volume is therefore not linear, and the percent area is larger in value than the percent volume for any ratio of radii between the two spheres. In general, the compressed breast and the dense tissue are not spherical. To investigate the empirical relationship between the percent area and the percent volume in the nonlinear situation, we applied least squares fits in several polynomial models to the data points in Fig. 7. The results are shown in Table II and Fig. 9. A comparison of Table I and Table II indicates that the $Y=kx^{2/3}$ model (x=percent fibroglandular tissue volume, Y=percent mammographic dense area) resulted in slightly larger rms residuals than the linear model. The model $Y=kx^m$ with m equal to 0.83 and 0.86, respectively, for CC- and MLO-views slightly reduced the rms residuals. The best fit was obtained from the model $Y=k_1x^m+k_2$. However, the situation that the percent projected area was negative when the percent volume was zero would not occur physically. Note that if the model was fitted to the percent area data segmented by MDEST (Fig. 8), the k_2 values would become positive, indicating that the nonzero k_2 values are likely caused by segmentation biases.

Overall, these models demonstrate that there is no simple mathematical relationship between the percent volume and the percent projected area but the values for the exponents appeared to be in a reasonable range. The relationship between the percent volumes of two 3-D objects, one within another, and their percent projected 2-D area depends on their shapes. For example, the closer the two volumes are to concentric cylinders of the same height, the closer the exponent is to unity. The spread of the data points can therefore be attributed to the various irregular shapes of the fibroglandular tissue in the breasts, the changes in the shapes of the fatty and fibroglandular tissue due to compression, as well as the uncertainties in the segmentation of both the mammograms and the MR images. Although the spread of the data points in the correlation plots is large, one can expect that when the mammographic density of a given patient is monitored over time, the variations in the projected dense area due to the geometric factors, described above, will actually be much less than that observed from the scatter plots among a large number of patients. In other words, the uncertainty in the estimated percent density from the serial mammograms of a given patient should be much less than those shown in

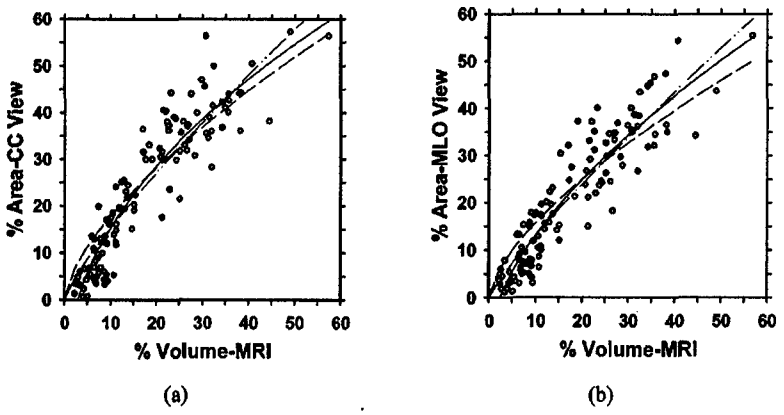


FIG. 9. Nonlinear fitting of the relationship between the percent volume and the percent area segmented by a radiologist with the least squares method. (a) CC view, (b) MLO view. Dashed line: $y=kx^{2/3}$; dashed–dotted–dotted line: $y=kx^m$; solid line: $y=k_1x^m+k_2$. The fitted parameters of the models, m, k, k_1 and k_2, are shown in Table II.

Fig. 7. The strong correlation observed between the percent dense area on mammograms and the percent volumetric fibroglandular tissue on MR images therefore indicates that a change in mammographic density can be a useful indicator of a change in percent fibroglandular tissue volume in the breast.

Recently, some researchers attempted to estimate the thickness of the fibroglandular tissue in local regions of the mammograms from the projected density.[37] This approach is expected to provide a more accurate estimation of the fibroglandular tissue volume if the true thicknesses of the fibroglandular tissue and fatty tissue can be determined at various locations of the projected breast region. The volume of the fibroglandular tissue can then be summed over the pixels in the breast region and the percent volume calculated. However, to obtain accurate measurements, this approach requires the knowledge of the sensitometric curve for the screen-film mammogram at the imaging facility (or use of a digital detector with linear response) and other physical parameters such as the scatter fraction, the beam quality and beam hardening, in addition to the compressed breast thickness and the breast shape profile at the periphery. Some of the requirements may be circumvented by using a look-up table predetermined with a phantom calibration. Other factors may have to be approximated or ignored, or require further corrections by imaging each mammogram with a calibration phantom placed adjacent to the breast. This method is still being developed and the accuracy of estimating the thickness of the local fibroglandular tissue from a mammogram is yet to be determined. To our knowledge, no study to date has demonstrated that fibroglandular tissue volume estimated from mammograms has a higher correlation with the percent volumetric fibroglandular tissue volume estimated from MR images or other volumetric methods than we found in our current study. Furthermore, even if the local fibroglandular tissue thickness on mammograms can be measured in a laboratory or in an academic center using elaborate calibration schemes, it is doubtful that these methods can be translated into routine clinical measurement in mammography clinics. Its use may then be limited to controlled clinical trials. An estimation of the percent dense area projected on mammograms is likely a more practical approach for breast density assessment. The high correlation between the percent dense area and the percent fibroglandular tissue volume on MR images as demonstrated in the current study further supports the validity of this approach.

V. CONCLUSION

In this study, we investigated the correlation between the percent mammographic dense area and the percent volumetric fibroglandular tissue as measured on MR images. A semiautomatic method was developed for segmentation of the MR images and a fully automated computerized method, MDEST, was used to segment the mammograms. The performance of MDEST on the set of mammograms used in this study was verified with an experienced radiologist's manual segmentation. The inter-observer variability in segmentation of MR images was found to be small with correlation coefficients of 0.99. The correlation between the percent volume on MR images and percent area segmented by a radiologist for either CC- view or MLO-view is 0.91. The correlation between percent volume and percent area estimated by MDEST is 0.91 and 0.89, respectively, for CC and MLO views. Mammographic density is thus highly correlated with the percent volumetric fibroglandular tissue in the breast. The high correlation indicates that changes in mammographic density may be a useful indicator of changes in fibroglandular tissue volume in the breast. Our computerized image analysis tool, MDEST, can provide a consistent and reproducible estimation of percent dense area on routine clinical mammograms. The automated image analysis tool may improve the sensitivity of quantifying mammographic density changes, thereby contributing to the understanding of the relationship of mammographic density to breast cancer risk, detection, and prognosis, and the prevention and treatment of breast cancer.

ACKNOWLEDGMENTS

This work is supported by U.S. Army Medical Research and Materiel Command Grants No. DAMD 17-01-1-0326, No. DAMD 17-02-1-0214, and No. DAMD 17-99-1-9294. The content of this paper does not necessarily reflect the position of the government and no official endorsement of any equipment and product of any companies mentioned should be inferred.

[a]Author to whom correspondence should be addressed. Jun Wei, Ph.D., Department of Radiology, University of Michigan, CGC B2103, 1500 E. Medical Center Drive, Ann Arbor, Michigan 48109. Phone: 734-647-8553; fax: 734-615-5513; electronic mail: jvwei@umich.edu

[1]A. F. Saftlas, R. N. Hoover, L. A. Brinton, M. Szklo, D. R. Olson, M. Salane, and J. N. Wolfe, "Mammographic densities and risk of breast cancer," Cancer (N.Y.) **67**, 2833–2838 (1991).

[2]N. F. Boyd, G. A. Lockwood, J. W. Byng, D. L. Tritchler, and M. J. Yaffe, "Mammographic densities and breast cancer risk," Cancer Epidemiology Biomarkers & Prevention **7**, 1133–1144 (1998).

[3]C. M. Vachon, C. C. Kuni, K. Anderson, V. E. Anderson, and T. A. Sellers, "Association of mammographically defined percent breast density with epidemiologic risk factors for breast cancer (United States)," Cancer Causes & Control **11**, 653–662 (2000).

[4]P. M. Krook, "Mammographic parenchymal patterns as risk indicators for incident cancer in a screening program: an extended analysis," Am. J. Roentgenolo. **131**, 1031–1035 (1978).

[5]R. L. Egan and R. C. Mosteller, "Breast cancer mammography patterns," Cancer (N.Y.) **40**, 2087–2090 (1977).

[6]B. Threatt, J. M. Norbeck, N. S. Ullman, R. Kummer, and P. Roselle, "Association between mammographic parenchymal pattern classification and incidence of breast cancer," Cancer (N.Y.) **45**, 2550–2556 (1980).

[7]M. Moskowitz, P. Gartside, and C. Mclaughlin, "Mammographic patterns as markers for high-risk benign breast disease and incident cancers," Radiology **134**, 293–295 (1980).

[8]I. Witt, H. S. Hansen, and S. Brunner, "The risk of developing breast cancer in relation to mammography findings," Eur. Radiol. **4**, 65–67 (1984).

[9]S. Ciatto and M. Zappa, "A prospective study of the value of mammographic pattern as indicators of breast cancer risk in a screening experience," Eur. Radiol. **17**, 122–125 (1993).

[10]E. Thurfjell, C. C. Hsieh, L. Lipworth, A. Ekbom, H. O. Adami, and D. Trichopoulos, "Breast size and mammographic pattern in relation to breast cancer risk," Eur. J. Cancer Prevention **5**, 37–41 (1996).

[11]I. Kato, C. Beinart, A. Bleich, S. Su, M. Kim, and P. G. Toniolo, "A nested case-control study of mammographic patterns, breast volume and

breast cancer (New York City, NY, United States)," Cancer Causes & Control **6**, 431–438 (1995).
[12] E. Sala, R. Warren, J. McCann, S. Duffy, N. Day, and R. Luben, "Mammographic parenchymal patterns and mode of detection: implications for the breast screening programme," J. Medical Screening **5**, 207–212 (1998).
[13] T. M. Salminen, I. E. Saarenmaa, M. M. Heikkila, and M. Hakama, "Is a dense mammographic parenchymal pattern a contraindication to hormonal replacement therapy?," Acta Oncol. **39**, 969–972 (2000).
[14] C. Byrne, C. Schairer, J. N. Wolfe, N. Parekh, M. Salane, L. A. Brinton, R. Hoover, and R. Haile, "Mammographic features and breast cancer risk: Effects with time, age, and menopause status," J. Natl. Cancer Inst. **87**, 1622–1629 (1995).
[15] N. F. Boyd, C. Greenberg, G. Lockwood, L. Little, L. Martin, J. Byng, Y. Martin, and D. Tritchler, "Effects at two years of a low-fat, high-carbohydrate diet on radiologic features of the breast: Results from a randomized trial," J. Natl. Cancer Inst. **89**, 466–467 (1997).
[16] D. V. Spicer, G. Ursin, Y. R. Parisky, J. G. Pearce, D. Shoupe, A. Pike, and M. C. Pike, "Changes in mammographic densities induced by a hormonal contraceptive designed to reduce breast cancer risk," J. Natl. Cancer Inst. **86**, 431–436 (1994).
[17] J. Brisson, R. Verreault, A. S. Morrison, D. Tennina, and F. Meyer, "Diet, mammographic features of breast tissue, and breast cancer risk," Am. J. Epidemiol. **130**, 14–24 (1989).
[18] J. N. Wolfe, "Mammography: Ducts as a sole indicator of breast carcinoma," Radiology **89**, 206–210 (1967).
[19] J. N. Wolfe, "The prominent duct pattern as an indicator of cancer risk," Oncology **23**, 149–158 (1969).
[20] J. N. Wolfe, "Breast patterns as an index of risk for developing breast cancer," Am. J. Roentgenol. **126**, 1130–1139 (1976).
[21] J. N. Wolfe, "Risk for breast cancer development determined by mammographic parenchymal pattern," Cancer (N.Y.) **37**, 2486–2492 (1976).
[22] I. E. Magnin, F. Cluzeau, C. L. Odet, and A. Bremond, "Mammographic texture analysis: An evaluation of risk for developing breast cancer," Opt. Eng. **25**, 780–784 (1986).
[23] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe, "Automated analysis of mammographic densities," Phys. Med. Biol. **41**, 909–923 (1996).
[24] J. W. Byng, N. F. Boyd, E. Fishell, R. A. Jong, and M. J. Yaffe, "The quantitative-analysis of mammographic densities," Phys. Med. Biol. **39**, 1629–1638 (1994).
[25] M. J. Yaffe, N. F. Boyd, J. W. Byng, R. A. Jong, R. Fishell, G. A. Lockwood, L. E. Little, and D. L. Tritchler, "Breast cancer risk and measured mammographic density," Eur. J. Cancer Prevention **7**, S47–S55 (1998).
[26] Z. Huo, M. L. Giger, D. E. Wolverton, and W. Zhong, "Computerized analysis of mammographic parenchymal patterns for breast cancer risk assessment: Feature selection," Med. Phys. **27**, 4–12 (2000).
[27] J. J. Heine and R. P. Velthuizen, "A statistical methodology for mammographic density detection," Med. Phys. **27**, 2644–2651 (2000).
[28] J. M. Boone, K. K. Lindfors, C. S. Veatty, and J. A. Seibert, "A breast density index for digital mammograms based on radiologists' ranking," J. Digit Imaging **11**, 101–115 (1998).
[29] J. N. Wolfe, A. F. Saftlas, and M. Salene, "Evaluation of mammographic densities: A case-control study," Am. J. Roentgenol., Radium Ther. Nucl. Med. **148**, 1087–1092 (1987).
[30] *American College of Radiology. Breast Imaging—Reporting and Data System (BI-RADS)*, 3rd ed. (American College of Radiology, Reston, VA, 1998).
[31] E. White, P. Velentgas, M. T. Mandelson, C. D. Lehman, J. G. Elmore, P. Porter, Y. Yasui, and S. H. Taplin, "Variation in mammographic breast density by time in menstrual cycle among women aged 40–49 years," J. Natl. Cancer Inst. **90**, 906–910 (1998).
[32] C. Zhou, H. P. Chan, N. Petrick, M. A. Helvie, M. M. Goodsitt, B. Sahiner, and L. M. Hadjiiski, "Computerized image analysis: Estimation of breast density on mammograms," Med. Phys. **28**, 1056–1069 (2001).
[33] S. J. Graham, M. J. Bronskill, J. W. Byng, M. J. Yaffe, and N. F. Boyd, "Quantitative correlation of breast tissue parameters using magnetic resonance and x-ray mammography," Br. J. Cancer **73**, 162–168 (1996).
[34] N. A. Lee, H. Rusinek, J. Weinreb, R. Chandra, H. Toth, C. Singer, and G. Newstead, "Fatty and fibroglandular tissue volumes in the breasts of women 20–83 years old: comparison of x-ray mammography and computer-assisted MR imaging," Am. J. Roentgenol. **168**, 501–506 (1997).
[35] B. Sahiner, N. Petrick, H. P. Chan, L. M. Hadjiiski, C. Paramagul, M. A. Helvie, and M. N. Gurcan, "Computer-aided characterization of mammographic masses: Accuracy of mass segmentation and its effects on characterization," IEEE Trans. Med. Imaging **20**, 1275–1284 (2001).
[36] B. Sahiner, H. P. Chan, N. Petrick, D. Wei, M. A. Helvie, D. D. Adler, and M. M. Goodsitt, "Classification of mass and normal breast tissue: A convolution neural network classifier with spatial domain and texture images," IEEE Trans. Med. Imaging **15**, 598–610 (1996).
[37] O. Pawluczyk, B. J. Augustine, M. J. Yaffe, D. Rico, J. Yang, G. E. Mawdsley, and N. F. Boyd, "A volumetric method for estimation of breast density on digitized screen-film mammograms," Med. Phys. **30**, 352–364 (2003).

An observer study comparing spot imaging regions selected by radiologists and a computer for an automated stereo spot mammography technique

Mitchell M. Goodsitt,[a] Heang-Ping Chan, Justin T. Lydick, Chaitanya R. Gandra, Nelson G. Chen, Mark A. Helvie, Janet E. Bailey, Marilyn A. Roubidoux, Chintana Paramagul, Caroline E. Blane, Berkman Sahiner, and Nicholas A. Petrick[b]
Department of Radiology, University of Michigan, Ann Arbor, Michigan 48109-0030

(Received 22 October 2003; revised 16 February 2004; accepted for publication 15 March 2004; published 27 May 2004)

We are developing an automated stereo spot mammography technique for improved imaging of suspicious dense regions within digital mammograms. The technique entails the acquisition of a full-field digital mammogram, automated detection of a suspicious dense region within that mammogram by a computer aided detection (CAD) program, and acquisition of a stereo pair of images with automated collimation to the suspicious region. The latter stereo spot image is obtained within seconds of the original full-field mammogram, without releasing the compression paddle. The spot image is viewed on a stereo video display. A critical element of this technique is the automated detection of suspicious regions for spot imaging. We performed an observer study to compare the suspicious regions selected by radiologists with those selected by a CAD program developed at the University of Michigan. True regions of interest (TROIs) were separately determined by one of the radiologists who reviewed the original mammograms, biopsy images, and histology results. We compared the radiologist and computer-selected regions of interest (ROIs) to the TROIs. Both the radiologists and the computer were allowed to select up to 3 regions in each of 200 images (mixture of 100 CC and 100 MLO views). We computed overlap indices (the overlap index is defined as the ratio of the area of intersection to the area of interest) to quantify the agreement between the selected regions in each image. The averages of the largest overlap indices per image for the 5 radiologist-to-computer comparisons were directly related to the average number of regions per image traced by the radiologists (about 50% for 1 region/image, 84% for 2 regions/image and 96% for 3 regions/image). The average of the overlap indices with all of the TROIs was 73% for CAD and 76.8%+/−10.0% for the radiologists. This study indicates that the CAD determined ROIs could potentially be useful for a screening technique that includes stereo spot mammography imaging. [DOI: 10.1118/1.1737492]

Key words: stereomammography, spot mammography, observer study

I. INTRODUCTION

It has been estimated that early detection could save the lives of about 30–50% of women who develop breast cancer[1,2] X-ray mammography is the most powerful screening tool we have for detecting breast cancer. However, it has limitations, especially in imaging breasts containing large dense tissue regions.[3–5] Optimal perceptibility in these regions may not be attained even with the new full-field digital mammographic systems due to the higher noise levels associated with these poorly penetrated regions and due to the inability to separate overlapping tissues. Alternative views and spot mammography are known to be beneficial in these situations. For example, Hayes *et al.* have reported that magnification and spot compression techniques improved mammographic specificity in 50.8% of the screening cases that were recalled for assessment at their center.[6] Specifically, sixteen 'equivocal' diagnoses became 'normal' or 'benign' and 15 of these patients avoided surgical biopsy. Twelve 'equivocal' diagnoses became 'malignant,' which helped surgical planning, and in all 12 cases, histology confirmed the diagnosis of malignancy." Hayes *et al.* concluded that special views are "necessary for the complete mammographic assessment of many screen-detected abnormalities." Furthermore, Faulk and Sickles found, in a study of the efficacy of spot compression-magnification and tangential views in the mammographic evaluation of palpable breast masses, that spot compression–magnification views depicted 97% of the masses; whereas, standard views depicted 87%.[7] Tangential views also depicted 87% of the masses with some of the masses that were detected in tangential views not being detected in the standard views and *vice versa.* In addition, they found that use of special views enabled radiologists to correctly predict benign or malignant status in 77% of cases as compared with correct prediction in 69% of cases with only standard views. It is important to note that the supplemental imaging techniques like spot and tangential views are only employed in diagnostic work-up studies. They are not employed at screening.

Spot imaging is performed either in contact or magnification mode. The radiologist first examines a contact mammogram, and identifies a suspicious region for spot imaging. The technologist then utilizes the identified location on the

mammogram to reposition the breast to her best ability such that the suspicious region is in the center of the field. She then compresses this region using a special, smaller spot compression paddle. The x-ray beam is also collimated to a smaller field size. The goal is to mechanically separate a suspicious lesion from the surrounding tissue for improved perception in the spot mammogram. One problem with this method is that the positioning of the breast involves some guesswork so the desired lesion may not necessarily be imaged at all, or it may not be imaged optimally.

We have been developing an automated version of the spot imaging technique. Our initial idea was to employ a computer aided detection (CAD) program to determine the locations of suspicious dense regions within a full-field digital mammogram, and to take a second digital mammogram of only those regions using automated collimation and automated spot compression along with a more penetrating exposure. This second separate "spot mammogram" would be taken within seconds of the full-field mammogram while the breast is maintained in the same position, but compressed more to improve tissue separation. We developed instrumentation to implement this method[8] including (1) a stretched Mylar membrane device that is placed between the large conventional paddle and the breast and acts to restrain the breast during the changeover to the smaller spot compression paddle, (2) an $x-y$ translator for positioning the spot paddle at the suspicious region, and (3) a secondary collimator that restricts the x-ray beam to the suspicious region area.[8] We also performed experiments with a compressible breast phantom that contained simulated masses. We found that when we spot-compressed a particular simulated mass in the phantom, it moved laterally out from underneath a dense overlying region and became visible in the spot mammogram. However, we were unable to reproduce the same result when we repeated the experiment multiple times. This was disappointing, but it made us aware that it can be difficult to position the spot paddle to produce the desired shearing force on a lesion, and even when the force is in the right direction, it may not be sufficient to move the lesion far enough out from overlying or underlying dense tissue to be seen. This prompted us to think of an alternative spot compression method—one that would be easier to implement, involve less equipment modification and produce better distinction between overlapping lesions.

The new technique is stereo spot digital mammography. It is very similar to the above technique except there is no changeover to or positioning of a spot paddle and no need for a breast restraining device to hold the breast in the same position while switching to the spot paddle. Instead, after the suspicious region is identified in the full-field image by the CAD program, a stereo pair of images is immediately acquired with the x-ray beam automatically collimated to image that region. The additional equipment required is a secondary collimator to restrict the x-ray beam to the desired region and a stereo workstation that would be located in the radiologist reading room. The x-ray tube or the focal spot would also have to shift by about 3 to 6 degrees in the left and right directions to generate the left- and right-eye images of the stereo pair. Stereoscopic image acquisition and display will enable radiologists to view the suspicious regions in three dimensions. This will reduce the tissue superposition problem inherent in conventional single projection mammography. Furthermore, in comparison with conventional spot compression, the automated technique should produce more accurate spot imaging of suspicious regions because it eliminates the need for the repositioning of the breast between the full-field and spot images, and the "spot" location is determined by a computer analysis of the digital full-breast image rather than estimated by eye from a radiograph. Since the method is automated and does not require the on-line review of a radiologist to determine the locations of the suspicious regions, it could potentially be used in screening and could potentially eliminate the need for diagnostic callback studies in many cases.

Another approach would be to perform full-field stereomammography instead of stereo spot mammography. This could be carried out either as a replacement for the conventional full-field mammogram or as a supplement to that mammogram. The advantages of limiting the stereo imaging to a spot region as opposed to full-field are that spot mode limits the amount of breast tissue exposed to additional radiation, and it decreases the volume of tissue that scatters radiation thereby improving image contrast. The anticipated dose for stereo spot mammography would be considerably less than that of a full-field mammogram because of the smaller field size. The exact dose for the spot technique would depend upon the size of the spot collimated region, the x-ray technique factors (kVp, target, filter, mAs), the x-ray beam half-value-layer, the tissue thickness, the breast composition (amount of glandular tissue) in the path of the spot-collimated x-ray beam, and the amount of glandular tissue that is exposed to scattered x-rays as well as the amount of the x-ray scatter. Stereo spot mammography does involve the aquisition of 2 spot images (the left- and right-eye images), instead of 1, so one might think that it would therefore require twice the dose of a single spot image. However, the eye–brain system integrates the noise from both images when they are viewed as a stereo pair, so theoretically, the same signal-to-noise ratio could be achieved by using only about half the dose for each image of the stereo pair. This is confirmed by the results of an experimental study recently published by Maidment *et al.*[9] on the effects of quantum noise and binocular summation on dose requirements for full-field stereoradiography. Maidment *et al.* found "the total dose needed to produce a stereoradiographic image pair is approximately 1.1 times the dose needed for a single projection in standard radiography."

Automated detection of suspicious dense regions is a crucial element of the stereo spot mammography method. In this paper, we describe an observer study that was performed to compare the suspicious regions selected by radiologists with those selected by a CAD program developed in our laboratory at the University of Michigan. We also compared these to true regions of interest (TROIs) containing the masses. At the time of our study, a substantial collection of full-field digital mammography (FFDM) images with biopsy results

was not available, and our CAD algorithms had not yet been adapted to FFDM images. Therefore, we employed digitized film images and our film-based CAD algorithms. Although the images and the CAD system are not completely representative of the images and analysis that will be employed in the eventual implementation of the automated stereo spot method, they provide valuable preliminary data about a new application of CAD, namely automated detection of ROIs to be worked-up with spot imaging.

II. MATERIALS AND METHODS

IRB approval was obtained to review 200 digitized mammograms for this study.

Five MQSA qualified radiologists participated as the readers. Their experience in reading mammograms ranged from 5.5 to 25 years (mean=13.7 years, standard deviation=7.6 years). The radiologists viewed digitized film mammograms that have been employed previously in the development of our CAD system. The film set included craniocaudal (CC) view and the mediolateral oblique (MLO) view mammograms of both breasts of patients at our clinics. The mammograms were digitized with a LUMISYS 85 laser film scanner (Lumisys, Inc., Sunnyvale, CA) at a pixel size of 50 μm $\times$50 μm. This digitizer has a gray level resolution of 12 bits and a nominal optical density (O.D.) range of 0 to 4. To keep the reading time reasonable (i.e., about 3 hours), we had the observers perform the study on 200 mammograms. All of the observers also repeated the study 3 to 13 months later for an evaluation of their reproducibility.

Each observer was given the task of outlining in each mammogram, 0 to 3 suspicious regions which in an ideal world they would have spot-imaged. A computer program, TRACEIMAGE, was developed for this study to allow the radiologists to trace the boundaries of the suspicious regions using a computer mouse. The TRACEIMAGE program incorporates a graphical user interface (GUI) that displays the digitized mammogram within a window on the computer monitor. The GUI includes slide bars for adjusting the contrast and brightness of the image, a display of a histogram of the grayscale values within the breast region in the image, a pull-down menu for selecting the pixel size of the displayed image (the choices are 200, 400 and 800 microns), slide bars for panning the image when the selected pixel size is 200 microns, a trace routine that permits the operator to outline the suspicious regions within the displayed mammogram via a series of computer mouse motions and mouse clicks, a button for erasing an individual trace within the image, and another button for erasing all traces within the image. The GUI was designed to be very user friendly. It automatically finishes a particular trace when a mouse click is within a preset number of pixels from the beginning point of the trace, and it then automatically proceeds to the next trace. Also, the user can re-adjust any of the traces in an image. Examples of the GUI display are shown in Figs. 1 and 2. The TRACEIMAGE program with its GUI was implemented on a PC computer with a high quality Hitachi (Hitachi, Ltd., Japan) Superscan 753 19 in. color monitor operating at a resolution of 1024$\times$768 in a noninterlaced true color mode with an 85 Hz refresh rate. The monitor was adjusted to meet DICOM standards, and the room lights were dimmed to a very low level during the observer studies. During the study, each observer analyzed each image separately without referring to previous or subsequent images. They were not told which MLO views corresponded with which CC views and *vice versa*. Furthermore, all patient identification information was removed from the images.

For the computer-selected regions, we employed a mass detection CAD program developed at the University of Michigan.[10] This program consists of 4 steps. First, the digitized mammogram is processed with a density weighted contrast enhancement (DWCE) filter that adaptively enhances local area contrast in order to emphasize mammographic structures. Second, an edge detection algorithm is employed to define the borders of the enhanced structures, resulting in a set of detected structures. Third, a local refinement algorithm, which includes erosion and K-means clustering, is applied to the detected structures to improve the accuracy of the borders and to split large connected regions. Fourth, the refined detected objects are classified as masses or normal breast structures based on the input of extracted morphological and textural features into a linear discriminant classifier. Potential masses are identified using decision thresholds that are based on the linear discriminant classifier score and the maximum number of marks allowed per image. For the present study, we adjusted the detection threshold of the CAD program[10] to mark between 0 and 3 regions in each image, with 3 the most likely number. In addition, rectangular bounding boxes that enclosed each known true mass in the set of mammograms were also obtained. These were regions identified by a radiologist from analyses of the mammograms along with associated pathology biopsy data and biopsy images. Of the 200 images that were evaluated by the radiologists in our observer study, 98 images contained TROIs. There were 83 images with a single true mass, 13 with two true masses, and 2 with three true masses. Thus, there was a total of 115 ($=83+13\times2+2\times3$) TROIs. Out of the 98 images with TROIs, 51 (52%) contained malignant masses. 75% of the exams with TROIs were worked–up with spot compression (57% of these were malignant and 43% were benign). It should be noted that our data set of 200 images was for 37 patients, and there were images from 2 or more years for 11 of the patients. Some of these patients had "normal" mammograms in earlier years and suspicious regions worked-up with spot mammography and/or lateral views in later years.

A second computer program, DENSECOMP, was developed to display sets of traced regions in each image. The traces for up to 3 readers can be displayed at once using different colors. For example, the traces for one radiologist would be filled-in as red, those for a second radiologist would be filled-in as green and those for a third radiologist would be filled-in as blue. Anywhere in the image where all three of the readers' traces intersected would be displayed as white. For the study described in this paper, we were interested in the intersections between the radiologist-selected regions of

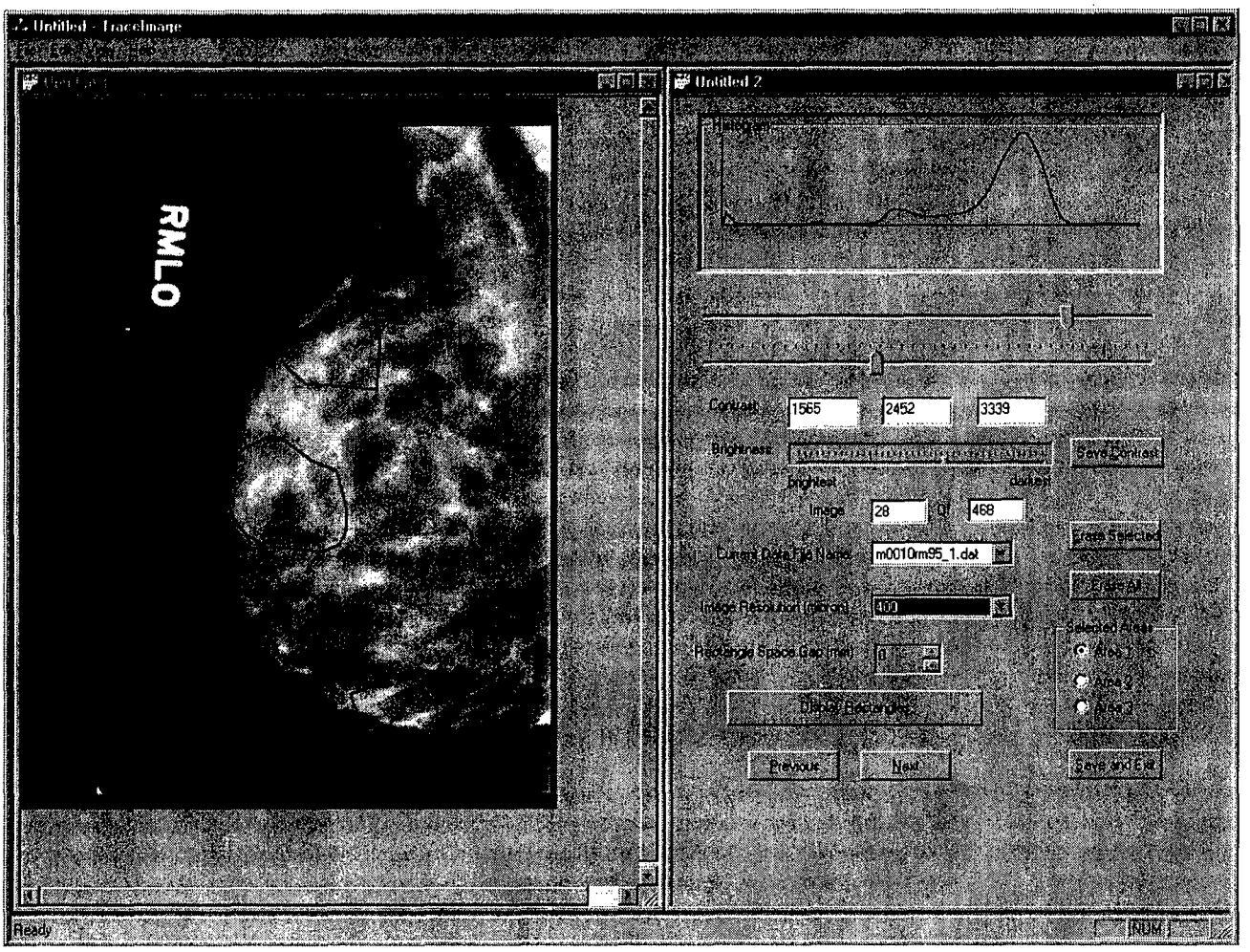


FIG. 1. Graphical user interface display showing one of the digitized mammograms in this study, 3 regions for spot imaging that were traced by a radiologist, a histogram of the pixel values within the breast region (top right), sliders (beneath histogram) for adjusting the range of pixel values that are mapped into an 8-bit output for display, and a pull-down menu for selecting the image resolution (the resolution displayed is 400 microns).

interest (ROIs) and those selected by the CAD program (for simplicity, the ROIs selected by the computer program will be referred to as CAD ROIs or CAD-selected ROIs in the following discussion) and also the intersections of both of these with the true mass regions. We found that the large number of possible intersections of the regions for 3 readers (e.g., 27 possible single intersections) can result in a complicated display that is difficult to interpret. We therefore decided to display the results for 2 readers at a time where a "reader" is either a radiologist, the CAD program or the true mass region. In addition to filling-in the ROIs with color, the program can also display the traces themselves superimposed on the mammogram in different colors. This allows one to see the suspicious lesion within the trace.

The DENSECOMP computer program computes overlap indices between the radiologist-selected ROIs and the CAD-selected ROIs. Furthermore, it computes overlap indices between the radiologists-selected ROIs and the TROIs and overlap indices between the CAD-selected ROIs and the TROIs. The overlap index for the radiologist vs CAD program comparison is defined as the area of the intersection between the radiologist-selected and CAD-selected ROIs divided by the area of the CAD ROI. This definition was chosen so that a value of 100% would be obtained if the CAD ROI was completely contained within the radiologist's ROI, indicating that the CAD ROI would be completely imaged with a spot region determined by the radiologist. Examples of 100% overlap are shown in Figs. 3(a) and 3(b). The overlap index between the radiologist's or CAD ROI and the TROI was defined as the area of intersection divided by the smaller area. This definition yields an overlap index of 100% if the radiologist's ROI or CAD ROI is completely contained within the TROI or if the TROI is completely contained within the CAD or radiologist's ROI. It is an indication of the degree to which there is a "hit" between the radiologist's or CAD ROI and the TROI. In addition, we computed the number of "hits" by determining the total number of times that the overlap indices with the TROIs were greater than or equal to a threshold value of 25%. The 25% value was considered to be a reasonable threshold for indicating agreement in ROIs. For the computation of the overlap indices between radiologist's and CAD ROIs, all possible pairings of the se-

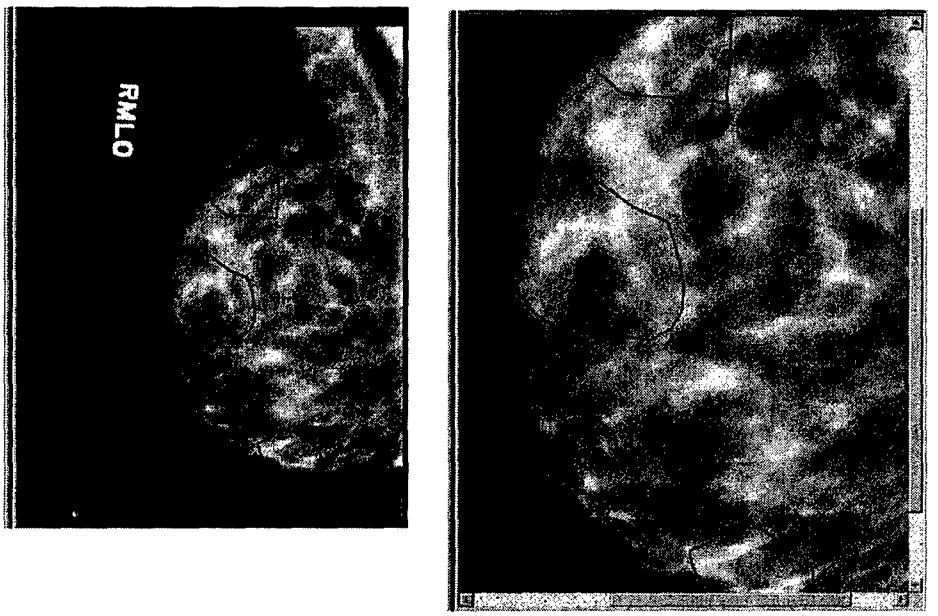


FIG. 2. An example of the GUI display of the same image at 400 micron (left) and 200 micron (right) resolution. Radiologists could use either display resolution for each image in the study, and could switch between the resolutions during their analysis. Many of the 200-micron resolution images were larger than the display window. The GUI included a panning feature to enable shifting of the viewed portion of the 200-micron image within the display window so the entire image could be viewed and analyzed.

lected regions were considered, and the indices were ordered from the largest to the smallest. For example, if the radiologist selected 3 regions and the CAD program selected 3 regions, there would be 9 possible intersections for the indices and the results would be listed in 9 columns with column 1 corresponding to the largest index. The radiologist's ROI-vs-TROI and CAD ROI-vs-TROI comparisons were special cases because we knew the true mass regions. For these particular comparisons, we computed the largest overlap indices individually for each TROI. For images with two or three TROIs, the largest overlap index for all of the TROIs was assigned to TROI #1, the largest overlap index for the remaining one or two TROIs was assigned to TROI #2, and if there was a third TROI, the largest overlap index for that TROI was assigned to TROI #3.

III. RESULTS

Comparisons of the ROIs selected by radiologists and the CAD program are shown in Fig. 3. In each image, the ROIs determined by the radiologist are filled-in in red, and those of the CAD program are filled-in in green. The intersections are displayed as yellow (red+green=yellow.) A comparison of the ROIs selected by each of the 5 radiologist readers [(a)–(e)] and the TROIs for an image containing 3 TROIs is shown in Fig. 4. The CAD selected ROIs are also compared with the TROIs in part (f) of this figure. Note that instead of filling-in the ROIs with color as in Fig. 3, the actual borders are displayed in black for the radiologist and CAD ROIs and in white for the TROIs in Fig. 4. The frequencies that each radiologist selected 0, 1, 2, and 3 ROIs in the 200 images at each reading session are listed in Table I, along with the frequencies for the CAD program.

The average values of the largest overlap indices between the radiologist-selected ROIs and the CAD-selected ROIs are listed for each radiologist and each reading session in Table II. A histogram summarizing the individual results for all of the radiologist vs CAD ROI pairings having the largest overlap index is shown in Fig. 5. The overall average overlap index is 69.6% with a standard deviation of 44.3%.

The average overlap indices between the TROIs and the

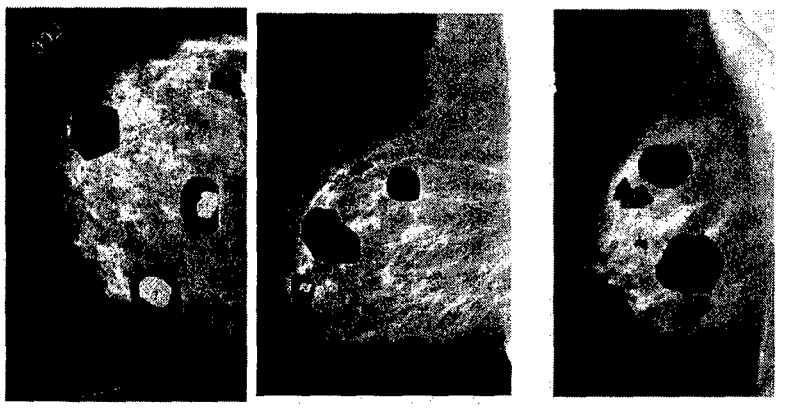


FIG. 3. Examples of the radiologist and computer selected ROIs. The radiologist ROIs are filled-in in red (black in figure), the CAD selected ROIs are filled-in in green (grey in figure), and the intersection areas are displayed in yellow (white in figure). In (a) the radiologist and computer agreed on 2 of the 3 ROIs; in (b) they agreed on one ROI; and in (c) they disagreed on all 3 ROIs.

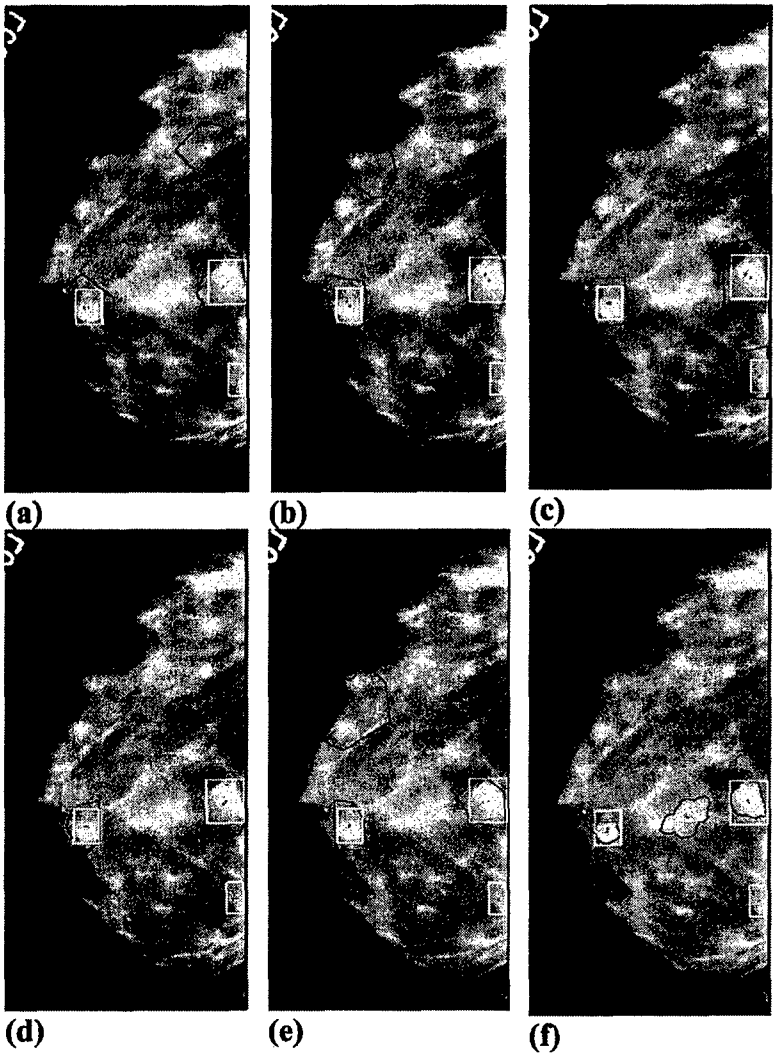


FIG. 4. Comparisons of the ROIs selected by the radiologists with the true mass regions (a)–(e) and the ROIs selected by the CAD program with true mass regions (f). The boundaries of the radiologist and CAD determined regions are traced in black, and the boundaries of the true regions are traced in white. Note that for this particular mammogram, radiologist (d) chose to select 2 regions instead of 3. Also, radiologist (c) was the only one whose selected regions intersected all 3 TROIs.

ROIs selected by the radiologists and the CAD program are listed in Table III. The average percentages of hits between each reader's selected ROIs and the TROIs are listed in Table IV.

IV. DISCUSSION

As shown in Table I, the radiologists executed the task of selecting up to 3 spot regions in each image differently, with some radiologists selecting almost 3 ROIs in each image, and others selecting far fewer with average numbers of ROIs per image even less than 1. The CAD program, as designed, selected 3 ROIs in nearly all (93%) of the images.

For the 5 radiologist-vs-CAD comparisons, the average agreement between the radiologist-selected ROIs for spot imaging and at least one of the CAD-selected ROIs for the entire set of images ranges from 43% to 98% (Table II). This unusually wide range of agreement can be explained if one compares the overlap indices listed in Table II with the average number of ROIs per image listed in Table I. A linear least squares fit between these data yields a correlation coefficient of 0.99 with a positive slope. That is, the greater the average number of ROIs per image selected by a radiologist, the greater the agreement with the computer. In general, the agreement or average overlap index between the radiologist-selected ROIs and at least one of the CAD-selected ROIs in each image was only about 50% for radiologists who traced an average of about 1 ROI per image. On the other hand, this agreement was much improved to 84% for radiologists who traced an average of about 2 ROIs per image, and it was about 96% for those who traced about 3 ROIs per image. Thus, it is anticipated that had we not given the radiologists full freedom to trace between 0 and 3 ROIs in each digitized mammogram for spot imaging, as in our original study design, but had we instead instructed them to trace 3 ROIs per mammogram similar to the number chosen for the computer, all of the radiologists would have had excellent agreement with the computer. Yet another factor that was not controlled in this observer experiment that could influence the agreement is the sizes of the radiologist selected ROIs, with greater sizes having greater probabilities of overlap with the CAD ROIs. The reproducibility of the overlap indices between the radiologist and CAD ROIs (Table II) for the two

TABLE I. Percentages of the 200 images in which each radiologist selected 0, 1, 2, and 3 ROIs during each reading session. Also, the corresponding percentages for the CAD program.

Reader	Reading session	% with 0 ROIs	% with 1 ROI	% with 2 ROIs	% with 3 ROIs	Average number of ROIs per image	Standard deviation
Radiologist a	1	1	0	16.5	87.5	2.86	0.43
Radiologist a	2	0	0.5	5.5	94	2.94	0.27
Radiologist b	1	1.5	12.5	38	48	2.32	0.75
Radiologist b	2	0.5	4.5	37	58	2.50	0.61
Radiologist c	1	31.5	53.5	13.5	1.5	0.85	0.70
Radiologist c	2	41	47.5	10	1.5	0.72	0.70
Radiologist d	1	39.5	50.5	10	0	0.70	0.64
Radiologist d	2	46.5	43	9	1.5	0.66	0.71
Radiologist e	1	2.5	23	40	34.5	2.06	0.82
Radiologist e	2	16.5	56	18	9.5	1.20	0.83
CAD program		0	2	5	93	2.91	0.35

TABLE II. Overlap indices between the ROIs selected by the radiologists and by the CAD computer program. The averages of the largest overlap indices for each image are listed for each radiologist in each reading session. These are averages for 200 images and up to 3 ROIs per image. The overlap index is defined to be the area of the intersection between the radiologist-selected and CAD-selected ROIs divided by the area of the CAD-selected ROI in percent.

	Radiologist a	Radiologist b	Radiologist c	Radiologist d	Radiologist e
Reading #1	94%	84%	51%	45%	82%
Reading #2	98%	86%	47%	43%	65%

TABLE III. Overlap indices between the ROIs selected by the readers and the true regions of interest (TROIs). The averages of the largest overlap indices for each TROI in each image are listed, as well as the averages for all TROIs. The overlap index in this case is defined to be the area of the intersection divided by the smaller of the TROI and reader ROI areas, in percent. There were 83 images with one TROI, 13 with two TROIs and 2 with three TROIs. Wilcoxon Signed Rank Test p-values relative to CAD for the All TROI results of each radiologist at each reading session are listed in the final column.

	TROI #1	TROI #2	TROI #3	All TROI	p-value
Radiologist a					
Reading #1	88%	58%	50%	84%[a]	0.003
Reading #2	94%	74%	46%	90%[a]	<0.0001
Radiologist b					
Reading #1	88%	70%	49%	85%[a]	0.008
Reading #2	88%	66%	46%	84%[a]	0.010
Radiologist c					
Reading #1	70%	42%	47%	66%	0.399
Reading #2	71%	44%	44%	67%	0.680
Radiologist d					
Reading #1	69%	45%	0%	65%	0.272
Reading #2	66%	48%	50%	64%	0.179
Radiologist e					
Reading #1	87%	50%	37%	81%	0.118
Reading #2	85%	67%	47%	82%[a]	0.018
Computer (CAD)	78%	53%	0%	73%	

[a]=significantly different from the CAD program at a 0.05 level or less.

TABLE IV. Average percentages of "hits." A hit is defined to occur whenever the overlap index between the reader ROI and the TROI is greater than or equal to 25%. Values relative to the total number of masses (TROIs) in the data set in percent are listed for each reading session. Wilcoxon Signed Rank Test p-values relative to CAD for each radiologist and reading session are also listed.

	Reading #1	p-value	Reading #2	p-value
Radiologist a	89%	0.074	93%[a]	0.006
Radiologist b	90%[a]	0.036	89%	0.062
Radiologist c	70%	0.081	70%	0.081
Radiologist d	69%	0.055	67%[a]	0.027
Radiologist e	87%	0.153	87%	0.135
CAD	80%			

[a]=significantly different from the CAD program at a 0.05 level or less.

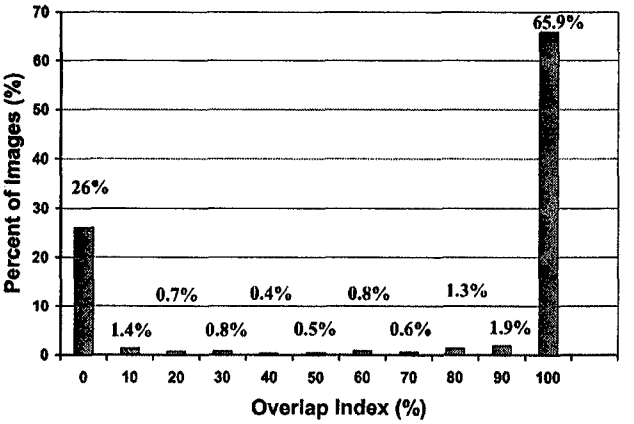


FIG. 5. A histogram summarizing the distribution of largest overlap indices between all radiologist and CAD-selected ROIs in 200 images including data for both reading sessions.

reading sessions is very good (within 4% for 3 radiologists and 17% for one) indicating each radiologist was very consistent in identifying suspicious ROIs for spot imaging at each session.

The agreement between the radiologist-selected ROIs and the true mass regions (TROIs) is very similar on average to that between the CAD-selected ROIs and the TROIs. For the entire set of 115 TROIs (Table III, column 4), the overall average overlap indices for the radiologists' ROI-vs-TROI comparisons ranged from 64% to 90% with a mean of 76.8%+/−10.0%, and the overall average overlap index for the CAD ROI-vs-TROI comparisons was 73%. As listed in Table III, 5 of the 10 radiologists' ROI-vs-TROI overlap indices were found to be statistically significantly different from the corresponding CAD ROI-vs-TROI overlap indices. Nonparametric statistics were employed for this comparison because, as can be deduced from Fig. 5, the data were not from a normal distribution. The Wilcoxon signed rank test was utilized. It is the nonparametric equivalent of the paired t-test. We found a very similar statistical result when we compared the ROI-vs-TROI overlap indices between pairs of radiologists. For example, for the 10 possible pairings of radiologists' results for the first reading sessions (the combination of 5 radiologists taken 2 at a time), we found 6 pairings were statistically significantly different. Thus, in terms of both the mean overlap indices of ROIs with TROIs and the statistical significance of the overlap indices, the CAD program performed comparably with an average radiologist.

The percentages of "hits" for which the overlap indices were greater than or equal to 25% for all TROIs ranged from 67% to 93% with a mean of 80.9%+/−10.6% for the radiologists' ROI-vs-TROI comparisons and was 80% for the CAD ROI-vs-TROI comparisons (Table IV). As shown in Table IV, 3 of the 10 (5 radiologists×2 reading sessions) numbers of hits with the TROIs for the radiologists were statistically significantly different from the number of hits for CAD. This is less than the 5 of 10 statistically significant differences for the radiologist-vs-CAD overlap indices discussed above, due to the thresholding effect associated with determining the number of hits. Overall the results for the number of hits confirm that the performance of the CAD method is very close to the average of the experienced radiologists. Finally, just as the agreement between the radiologist- and CAD-selected ROIs was highly correlated with the average number of radiologist-selected ROIs per image, so too was the number of "hits" highly correlated with this average number ($r=0.91$). That is, radiologists who selected more ROIs per image had a higher probability of hits with the TROIs.

The values of the radiologists' ROI–vs-TROI overlap indices were nearly identical for the two reading sessions [reproducibility within 1–7% for all TROIs (Table III, column 4) as were the percentages of "hits" (within 4%, Table IV]. This indicates the radiologists were very consistent in their selections of suspicious regions for spot imaging relative to the TROIs.

It should be acknowledged that the Hitachi monitor we employed for image display might have influenced the observer study results. The grayscale contrast, brightness and resolution of this monitor are inferior to those of a 2000 line physician's read monitor. Also, we did not employ full (50 μ)-resolution mammograms in this study. Nevertheless, the radiologists all felt the display of the 200 μ and 400 μ resolution images on the Hitachi monitor was adequate for the task of identifying suspicious density and mass regions.

The good agreement between the CAD-selected ROIs and the TROIs indicate that the CAD mass detection program has promise in an implementation of automated stereo spot mammographic imaging of dense areas. We have also developed a microcalcification detection program for CAD. The evaluation of the CAD microcalcification program for automated stereo spot mammographic imaging will be pursued in future studies.

Several practical issues would have to be addressed for the eventual implementation of the automated stereo spot technique including: the number of acceptable false positives for the CAD program, the minimum and maximum sizes of the spot areas, and criteria for determining whether to combine 2 or more ROIs into one. A secondary collimator would have to be designed and built to restrict the x-ray beam to the ROI region at the two projections involved in stereo spot

imaging. In practice, the positions of the collimator blades would be determined from ROI (s) obtained from CAD evaluation of the full-field mammogram, taking into account calibration and geometric factors. The blades would automatically collimate the beam as the mammography technologist shifted the x-ray tube first to the left (or right) and then to the right (or left) to take the stereo spot pair of images. Or, ideally, the x-ray tube or the focal spot should be shifted to the left and the right positions automatically to minimize the time between the images in the stereo pair and thus the breast compression time. The full-field mammograms along with any stereo spot images would be sent to a physician's workstation for interpretation by the radiologist. That workstation would incorporate stereo image display capability. This can be accomplished by adding a stereo graphics board and a transmitter for synchronizing liquid crystal glasses such that the radiologist's left eye would only see the left-eye spot image and the radiologist's right eye would only see the right-eye spot image. We have employed such systems in our stereo mammography research.[11–15]

A potential limitation of the automated stereo spot mammography method is that radiologists' abilities to fuse images and see them in stereo can be variable. That is, a certain percentage of radiologists may have poor binocular or stereo acuity, which unlike monocular visual acuity cannot be corrected with glasses. However, stereo acuity may be improved through training and use of depth cues provided by three-dimensional 3-D pointers or 3-D cursors. We have developed such cursors[11–13] and may investigate this aspect of stereo imaging in the future.

An alternative method that is not limited by the stereo acuity of the observer and yet should solve the tissue superposition problem would be to perform digital tomosynthesis[16–18] of the spot regions. The equipment required for such an automated spot tomosynthesis method is considerably more expensive than that required for stereo spot imaging. A potential advantage of spot tomosynthesis over full-field tomosynthesis is that the analysis could be concentrated on the suspicious areas. However, CAD techniques could also be employed on the full-field tomosynthesis images to highlight suspicious areas and speed up image analysis, obviating the need for spot tomosynthesis. Whether spot tomosynthesis has utility or not will be determined as we gain more experience with this imaging method.

V. SUMMARY AND CONCLUSIONS

Our observer study showed the regions identified by radiologists as being suspicious and warranting workup spot imaging are in good agreement with those identified by our CAD program. Also, both the radiologist- and the CAD-selected ROIs are in good agreement with the true mass regions as determined from analysis of mammograms and biopsy pathology results. Thus CAD could be used to identify suspicious regions for automatic stereo spot imaging. The stereo spot images are expected to improve perception of lesions that are camouflaged by overlying and underlying tissues in conventional mammograms. Stereo spot imaging may permit the discrimination of pseudo masses produced by the superposition of overlapping tissues in conventional mammograms, and it may enable better appreciation of the 3-D characteristics of lesions and microcalcifications. Thus, stereo spot imaging has the potential to improve the sensitivity and specificity of mammography. Since the method can be automated, and does not require the presence of on-site radiologists, it can be used in screening. A potential pitfall is the variable ability amongst radiologists to view images stereoscopically. This might be solved through stereo vision training. Another possibility would be to employ spot tomosynthesis of the suspicious regions.

ACKNOWLEDGMENTS

This work is supported by U. S. Army Medical Research and Materiel Command Grants No. DAMD17-99-1-9294 and No. DAMD17-98-1-8210. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred.

[a]Corresponding author: Mitchell M. Goodsitt, Ph.D., Department of Radiology, University of Michigan Hospitals, Room B1 F510C, 1500 East Medical Center Drive, Ann Arbor, Michigan 48109-0030. Office: 734-936-7474; fax: 734-936-7948; electronic mail: goodsitt@umich.edu
[b]Also at CDRH, FDA, HFZ-142, 12720 Twinbrook Parkway, Rockville, Maryland 20857.

[1] S. Shapiro, "The status of breast cancer screening; a quarter century of research," World J. Surg. **13**, 9–18 (1989).
[2] M. S. O'Malley, S. W. Fletcher, and B. Morrison, "Does screening for breast cancer save lives?" in *Preventing Disease*, edited by R. S. Laurence and R. Goldbloom (Springer-Verlag, New York, 1989), pp. 251–364.
[3] R. E. Bird, T. W. Wallace, and B. C. Yankaskas, "Analysis of cancers missed at screening mammography," Radiology **184**, 613–617 (1992).
[4] L. D'Angincourt, "Technique is everything when breast is dense," Diagn. Imaging **10**, 57–61 (Sept. 1993).
[5] V. P. Jackson, R. E. Hendrick, S. A. Feig, and D. B. Kopans, "Imaging of the radiographically dense breast," Radiology **188**, 297–301 (1993).
[6] R. Hayes, M. Michell, and H. B. Nunnerly, "Evaluation of magnification and paddle compression techniques in the assessment of mammographic screening detected abnormalities," Clin. Radiol. **44**, 158–160 (1991).
[7] R. M. Faulk and E. A. Sickles, "Efficacy of spot compression magnification and tangential views in mammographic evaluation of palpable breast masses," Radiology **185**, 87–90 (1992).
[8] M. M. Goodsitt, H. P. Chan, H. Huang, and C. Zhou, "Automated spot mammography for improved imaging of dense breasts," Radiology **217**, 346 (2000).
[9] A. D. A. Maidment, P. R. Bakic, and M. Albert, "Effects of quantum noise and binocular summation on dose requirements in stereoradiography," Med. Phys. **30**, 3061–3071 (2003).
[10] N. Petrick, B. Sahiner, H. P. Chan, M. A. Helvie, S. Paquerault, and L. M. Hadjiiski, "Breast cancer detection: Evaluation of a mass detection algorithm for computer-aided diagnosis: Experience in 263 patients," Radiology **224**, 217–224 (2002).
[11] M. M. Goodsitt, H. P. Chan, and L. Hadjiiski, "Stereomammography: Evaluation of depth perception using a virtual 3D cursor," Med. Phys. **27**, 1305–1310 (2000).
[12] M. M. Goodsitt, H. P. Chan, K. L. Darner, and L. M. Hadjiiski, "The effects of stereo shift angle, geometric magnification, and display zoom on depth measurements in digital stereomammography," Med. Phys. **29**, 2725–2734 (2002).
[13] M. M. Goodsitt, H. P. Chan, J. M. Sullivan, K. L. Darner, and L. M. Hadjiiski, "Evaluation of the effect of virtual cursor shape on depth measurements in digital stereomammograms." *Proceedings of the 5th International Workshop on Digital Mammography*, Toronto, Canada, 11–14,

June 2000; IWDM 2000, *5th International Workshop on Digital Mammography*, edited by M. Yaffe (Medical Physics Publishing, Madison, WI, 2001), pp. 45–50.

[14] H. P. Chan, M. M. Goodsitt, K. L. Darner, J. M. Sullivan, L. M. Hadjiiski, N. Petrick, and B. Sahiner, "Effects of stereoscopic imaging technique on depth discrimination," in Ref. 13, pp. 13–18.

[15] H. P. Chan, M. M. Goodsitt, L. M. Hadjiiski, J. E. Bailey, K. Klein, K. L. Darner, and B. Sahiner, "Effects of magnification and zooming on depth perception in digital stereomammography: An observer performance study," Phys. Med. Biol. **48**, 3721–3734 (2003).

[16] L. T. Niklason *et al.*, "Digital tomosynthesis in breast imaging," Radiology **205**, 399–406 (1997).

[17] R. L. Webber, H. R. Underhill, and R. I. Freimanis, "A controlled evaluation of tuned-aperture computed tomography applied to digital spot mammography," J. Digit Imaging **13**, 90–97 (2000).

[18] S. Suryanarayanan, A. Karellas, S. Vedantham, S. P. Baker, S. J. Glick, C. J. D'Orsi, and R. L. Webber, "Evaluation of linear and nonlinear tomosynthetic reconstruction methods in digital mammography," Acad. Radiol. **8**, 219–224 (2001).

To be published in the Proceedings of the International Workshop on Digital Mammography 2004

DEVELOPMENT OF AN ADD-ON ASYMMETRIC COLLIMATOR FOR AUTOMATED STEREO-SPOT DIGITAL MAMMOGRAPHY

Mitchell M. Goodsitt, PhD[1], Heang-Ping Chan, PhD [1], Justin T. Lydick[1], Dennis Kayner, BS[2], and Mark A. Helvie, MD[1]

Departments of Radiology[1] and Orthopedic Surgery[2], University of Michigan, Ann Arbor, MI

1.1 Introduction

Despite its wider dynamic range and improved contrast resolution compared to screen-film, digital mammography has not eliminated the need for diagnostic workups using alternative views and spot compression. Such imaging is needed when there is poor conspicuity in dense breast regions, and when there is tissue overlap. The advantages of these diagnostic imaging techniques for screen-film are well known. For example, Hayes et al (Hayes et al, 1991) found that magnification and spot compression techniques improved mammographic specificity in 50.8% of their diagnostic workup cases, and Faulk and Sickles (Faulk and Sickles, 1992) found perception of masses increased from 87% for standard views to 97% with spot compression magnification. Similar improvements would be expected for digital mammography.

Spot imaging requires that a radiologist be present to review the standard projection mammograms and identify suspicious regions for further workup. As such, it cannot be used during screening. We have been developing an automated spot imaging technique that could be used at screening.

2.1 First Automated Spot Compression/Collimation System

The initial method we proposed and pursued employed computer aided detection (CAD) rather than a radiologist to identify suspicious regions, and automated collimation and automated

application of spot compression to those regions. Some of the instrumentation that University of Michigan Mechanical Engineering students built for our first generation automated spot imaging device are shown in figure 1.

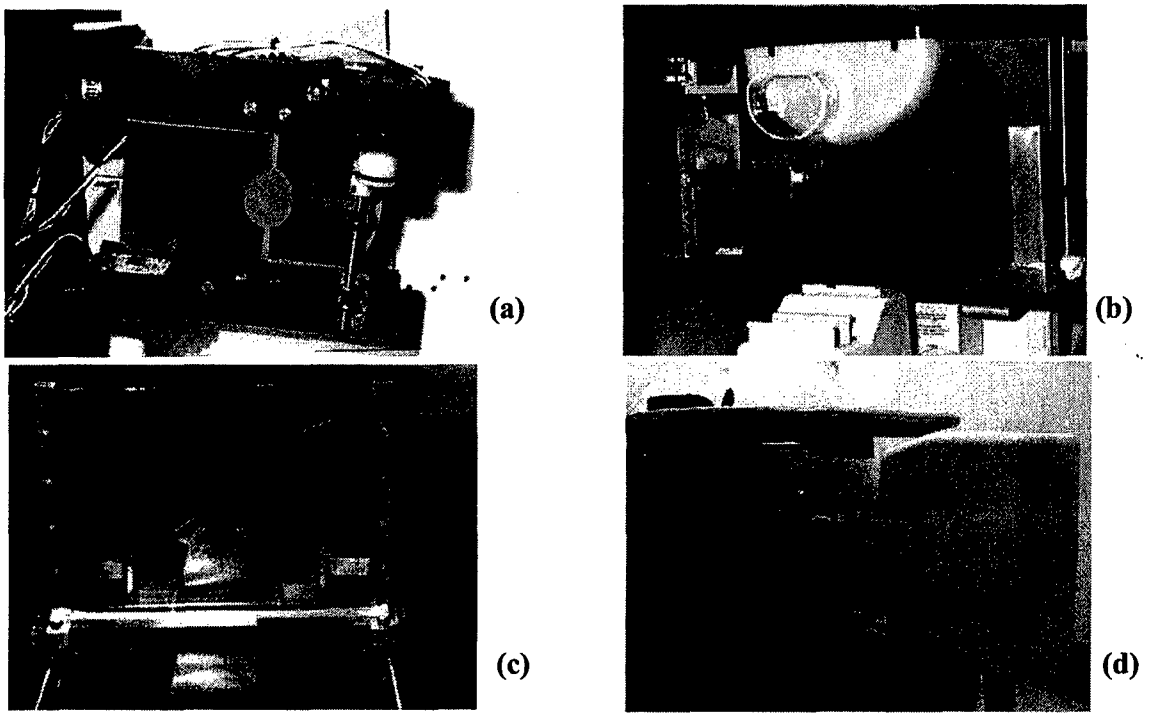


Figure 1. First Generation Automated Spot Compression Instrumentation:
a) spot collimator with 2 stepper motor drives, one to rotate the collimator and the other to change the size of the collimator opening, b) x-y translator for manually positioning the spot compression paddle at the location determined by the CAD program, c) breast restraining device that keeps the breast in a fixed position during the changeover from the standard compression paddle to the spot compression paddle, d) spot compression device attached to our research Fischer breast biopsy unit which employs a CR plate as the digital detector.

The initial system was mounted to a research breast biopsy unit that used a computed radiography (CR) digital detector. This was before our department purchased a flat panel digital mammography system (GE Senographe 2000D) that we have subsequently used for our automated spot compression research. The imaging procedure with our initial design applied to a full-field digital mammography unit (with a rapid readout detector) would entail inserting a thin membrane x-ray transmissive retraining device similar to that in figure 1(c) between the

standard compression paddle and the breast, the acquisition of a standard digital mammogram, CAD analysis of that mammogram to identify a suspicious region for spot imaging, retraction of the standard compression paddle, insertion of the spot compression paddle that would be moved with the x-y translator to compress the breast at the region of interest, automated supplemental collimation of the x-ray beam to the region of interest, and acquisition of the spot compressed image. The translation device for the positioning of the spot collimator could be automated with the use of stepper motors.

A key advantage of the automated spot concept as compared with conventional spot mammography is that it eliminates errors associated with the technologist judging by eye from the original mammogram where to re-position the breast such that the region identified by the radiologist is beneath the spot paddle. The idea behind spot compression is that confining the compression to a small area can result in a shearing force that should help separate overlapping tissues and lesions. We performed spot compression experiments on custom-made compressible breast simulating phantoms (made for us by CIRS, Inc., Norfolk, VA) and found that this was true some of the time. However, the spot compression could also sometimes squeeze overlapping lesions closer together making them more difficult to distinguish. This and the relative complexity of the instrumention required for our initial design prompted us to think of an alternative spot imaging technique – automated stereo spot mammography.

3.1 Stereo Spot Mammography

Stereo mammography produces a 3D image that improves the conspicuity of overlapping tissues. In a study comparing full field digital stereo mammography to standard film mammography,

Getty et al (Getty et al, 2001) found that diagnostic accuracy increased from 0.83 for film to 0.86 for digital stereo mammography, which was statistically significant. They also noted that with stereo mammography, radiologists were able to detect 39 new lesions in 129 cases. Since the eye-brain system integrates the noise from each image in a stereo pair, the total radiation dose requirement for digital stereo mammography is approximately equal to that of a standard single projection mammogram. This has been confirmed in an experimental study by Maidment et al. (Maidment et al, 2003) While full-field digital stereo mammography could be used for workup imaging during screening, restricting the field to a spot region has several advantages including reduced dose (since only the spot area is exposed rather than the entire breast), improved contrast (since a smaller scattering volume is exposed), and potentially less false positives (since the radiologist reviews stereo images of a restricted region containing only the lesions of interest rather than reviewing the entire breast in 3D).

3.1.1 Stereo Spot Procedure

The stereo spot imaging procedure that we foresee begins with the acquisition of a full-field digital mammogram. Next, within seconds, a CAD program is applied to that image, and it determines whether there are any suspicious regions that would require spot imaging. If there is a suspicious region, the coordinates of the region are fed directly into a program that drives stepper motors for positioning the blades of an automated spot collimator The technologist would shift the x-ray tube to a pre-determined angle, left of the central ray, the spot collimator blades would automatically adjust and the left-eye image of the stereo spot pair would be acquired. Next, the technologist would shift the x-ray tube to the same pre-determined angle to the right of the central ray, the collimator blades would automatically adjust, and the right-eye

image of the stereo spot pair would be acquired. The compression paddle would then be automatically released. All images would be transferred to a special workstation that includes a stereo graphics board, a high-quality, fast refresh rate (e.g. 120 Hz) display monitor, and stereo glasses with LCD shutters that are synchronized with the stereo graphics board such that the viewer's left eye only sees the left-eye image, and the viewer's right eye only sees the right-eye image.

3.1.1.1 Stereo Spot System

The system we have been developing includes a PC and software programs that we wrote to detect suspicious ROIs for spot imaging, operate the automated collimator, and control the stereo display. Our stereo display/viewing system consists of a Dome Model Md8/PCI Stereographics board, a Megascan MD8-4820-LS 2k Stereo monitor (120 Hz refresh rate, 2304 x 1800 stereo), StereoGraphics CrystalEyes infra-red emitter and LCD glasses, and an in-house developed graphical user interface that includes features such as window level and width adjustment, switching assignment of left and right-eye images so the stereo image appears to extend either into the monitor or out towards the viewer, display of mono and stereo images, electronic shutters, zoom, stereo cursors, and variable adjustment of the separation (parallax) between the left- and right-eye images. We employ a mass detection CAD program (Petrick et al, 1999; Petrick et al 2002) for the identification of the ROIs. This program allows the user to adjust the detection threshold to mark at most a specified number of regions in the image. We recently published the results of a study comparing the spot ROIs selected by this CAD program and the spot ROIs selected by 5 MQSA qualified radiologists. (Goodsitt et al, 2004) We also compared the ROIs selected by the CAD program and the radiologists to "Ground True ROIs" – regions

identified by a radiologist based on the mammograms, pathology biopsy data, and biopsy images. In general, we found that the performance of the CAD program was about equal to the average performance of the 5 radiologists. Therefore, the CAD program appears promising for the automated stereo spot imaging task.

4.1 Stereo Spot Collimator

The spot collimator that we developed was custom designed to attach to a GE Senographe 2000D full field digital mammography (FFDM) system. Since this is a clinical mammography system, the spot collimator could not be attached permanently or even semi-permanently. Therefore, one of the design requirements was quick and easy attachment and detachment of the research spot collimator. Once attached, it was required that the spot blades open far enough so that they are not be visible in the original full-field digital mammogram. Another important design requirement was that the spot collimator's outer dimensions not extend significantly beyond those of the present mammography unit gantry casing so that the collimator did not reduce patient clearance for imaging in the various projections. It was also desired that the collimator be capable of restricting the x-ray beam to rectangular areas that are off-center (asymmetric) and that the blades be capable of moving sufficiently far so as to collimate the x-ray beam to ROIs at almost any location within the breast.

To satisfy the goal of easy attachment, the collimator was designed to slide into the side slots for the face shield on the GE 2000D FFDM unit. The outer edges of the collimator follow the contour of the gantry casing, as shown in figure 2, below.

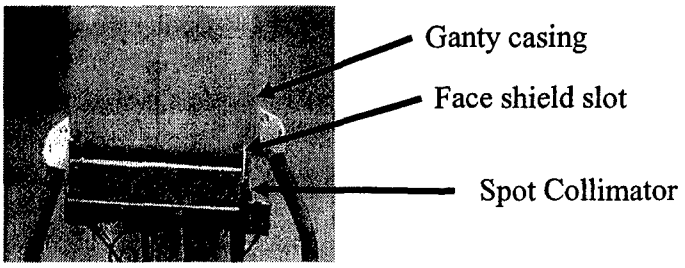


Figure 2. Spot collimator attached to GE Senographe 2000D FFDM system.

The complete collimator and the individual sections responsible for collimation in the lateral and chest-wall – anterior directions are shown in figure 3. The collimator was built by one of the co-authors of this paper, Dennis Kayner, who is a machinist in the Orthopedic Surgery Department.

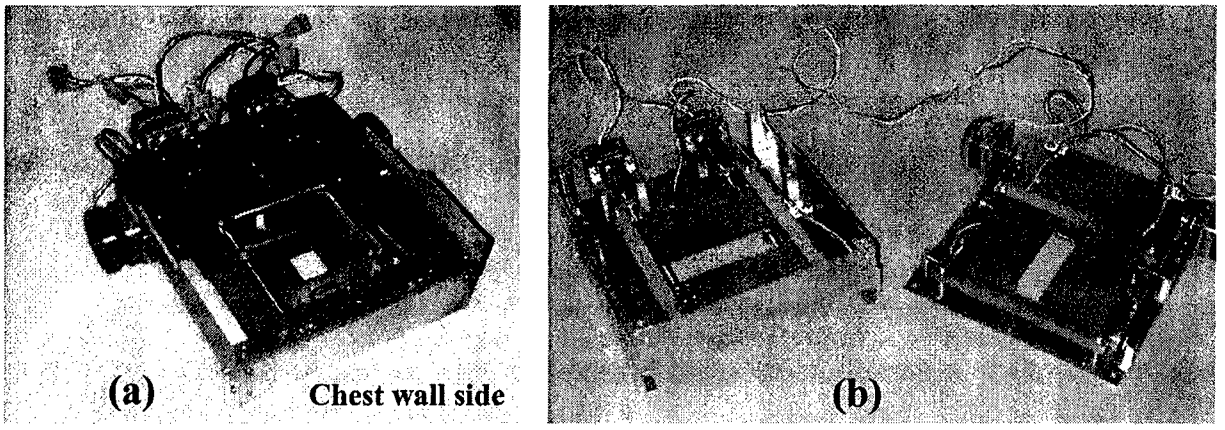


Figure 3. (a) Complete spot collimator. (b) Individual collimator sections

The collimator blades are made of 2-mm-thick steel, which is the same material and thickness as the blades used in the GE Senographe 2000D collimator. There are 4 stepper motors in the spot collimator. They are located towards the rear so as to not interfere with the patient's face in the CC-view. Two of the motors directly drive the anterior and chest wall collimator blades, and two are attached to belts for driving the left and right lateral blades. The specific motors employed are Shihno Kenshi (Japan) SST39D1010, which rotate 8 degrees per step. The stepper controller is a TMG (The Motion Group, Clovis, CA) MMC-4S. This system is very versatile.

It includes optical home sensors, it can trace true point-to-point patterns, and it utilizes a simple command set that permits the use of any programming language including Visual C to generate motion control sequences.

In-house software was written to allow a user (or a CAD program) to identify a rectangular spot region of interest within a full-field digital mammogram, and to collimate the beam to the region for stereo spot image acquisition as described below. Software was also written to calibrate the motion of the collimator blades. By moving the blades known numbers of steps and measuring the corresponding pixels moved in the projected image of the blades in the digital mammograms, we established that there were about 5.55 steps per (100 μ) pixel for all 4 blades. We determined the linear relationships between the number of steps traversed from each blade's home position and the x, y pixel coordinates of the projected collimator edge in the digital image for the x-ray tube located in the non-stereo orientation (0 °) and in the stereo orientations (6° left and 6 ° right.)

In order to ensure that a mass located within the specified ROI is imaged entirely in the stereo mode, the geometry of the imaging situation must be considered. Figure 4 illustrates the projection geometry for stereo imaging. The cross-hatched region in this figure is the area that projects to form the ROI on the standard mammogram. To guarantee that any object within this region is imaged in stereo, the spot collimator blades must be positioned so that the projections of their innermost edges correspond with the XLS and XRS image coordinates shown in the figure. As shown in the rightmost section of the figure, the x-ray tube rotates about a fulcrum. Because the stereo angle is small (typically 3 to 6°), the x-ray tube shift can be approximated as a displacement in the horizontal direction, and it can be shown that

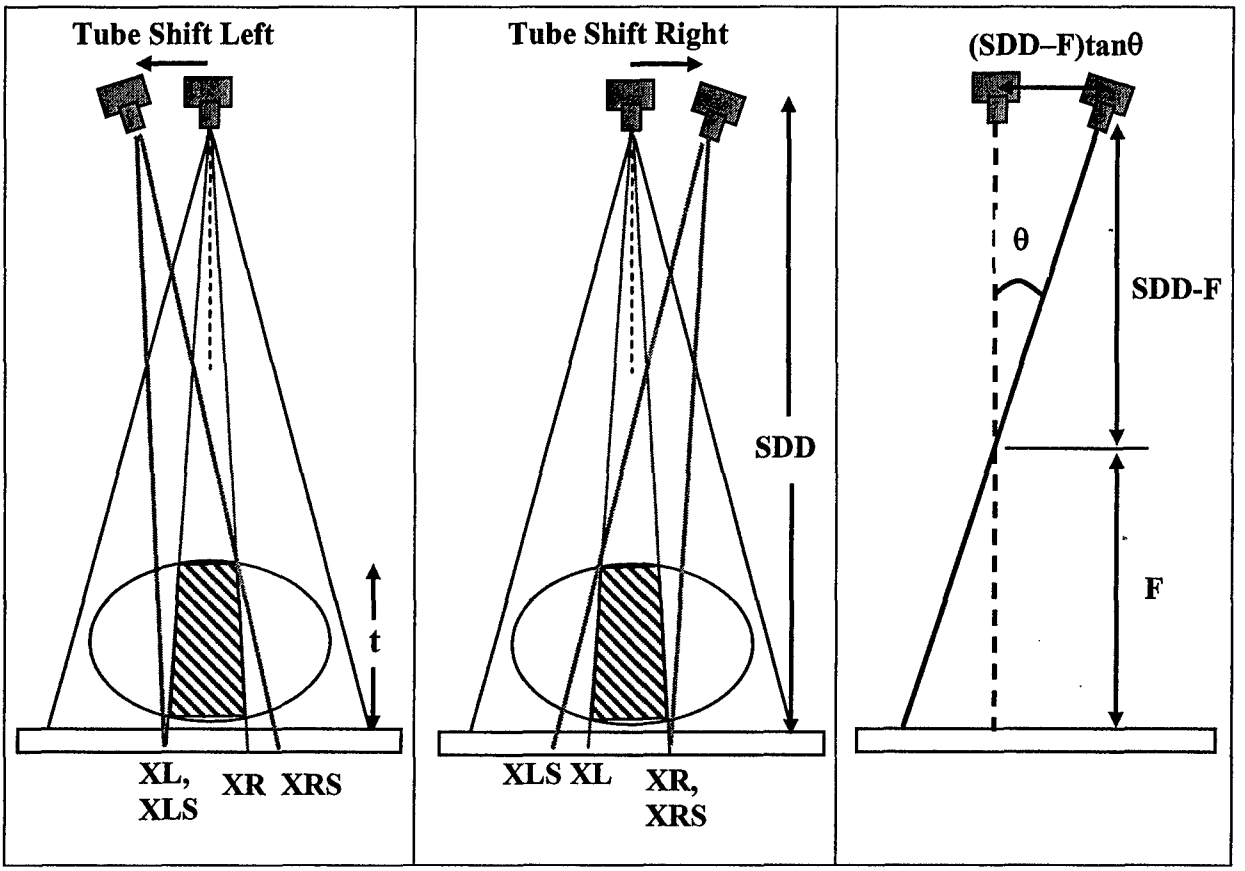


Figure 4. Geometry for Stereo Image Acquisition.

the XLS and XRS coordinates for left-shift image are given by the equations:

$$XLS = XL$$

$$XRS = XR + t * \tan\theta * \left(\frac{SDD - F}{SDD - t}\right)$$

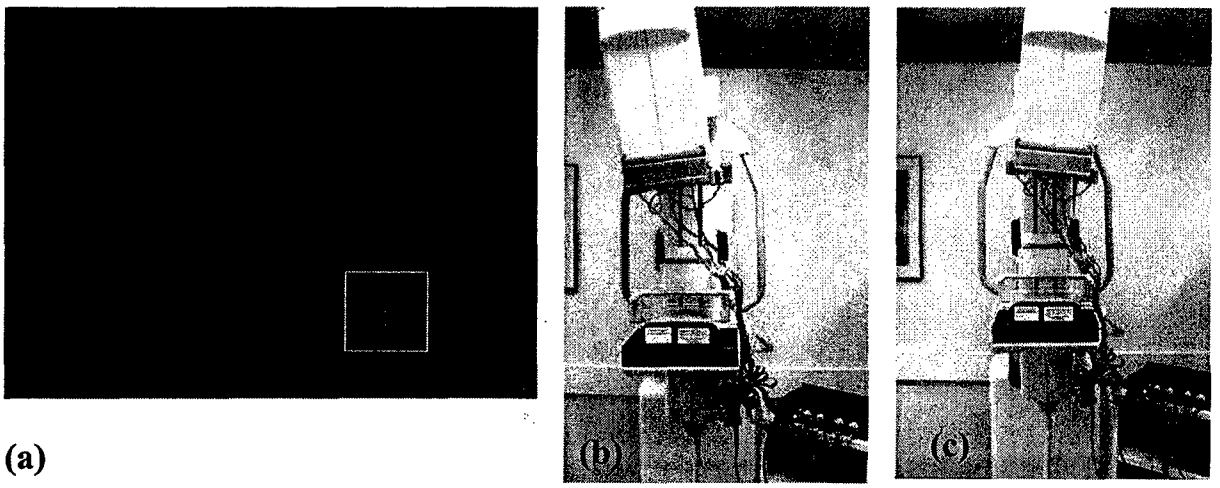
The corresponding coordinates for the right-shift image are:

$$XLS = XL - t * \tan\theta * \left(\frac{SDD - F}{SDD - t}\right)$$

$$XRS = XR$$

where XL and XR are the transverse coordinates of the ROI for the 0° (non-stereo) image, SDD is the source-to-detector distance, θ is the stereo-shift angle, F is the fulcrum-to-detector distance, and t is the thickness of the breast. The Y-coordinates of the collimator blades and ROI in the chest wall-anterior direction do not change with the stereo shift.

5.1 Experimental Study We performed an experiment in which we took full-field and stereo spot collimated images of a custom-made stereoscopic breast phantom (CIRS, Inc.) containing overlapping masses. The experimental setup and full-field image with a marked ROI are shown in figure 5. 0° spot, and 6° left-shift and 6° right-shift spot images are shown in figure 6.



(a)

Figure 5. (a) Full field digital mammogram showing user-selected rectangular ROI for spot imaging. Setup for left (b) and right (c) shift stereo spot acquisition.

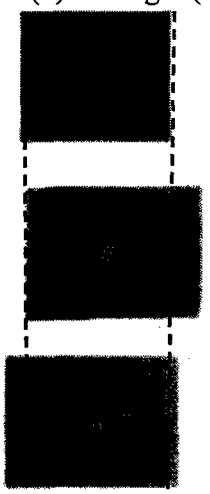


Figure 6. Top = 0° spot image, middle = 6° left-shift spot image, bottom = 6° right-shift spot image

6.1 Conclusion In conclusion, an automated stereo spot digital mammography system has been developed, and initial results are promising.

Acknowledgments

This work was supported by U. S. Army Medical Research and Materiel Command Grants DAMD 17-99-1-9294 and DAMD 17-98-1-8210. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred.

References

Faulk R.M. and E.A. Sickles. 1992. Efficacy of spot compression magnification and tangential views in mammographic evaluation of palpable breast masses. Radiology 185: 87-90.

Getty D.J., R.M. Pickett, and C.J. D'Orsi. 2001. "Stereoscopic digital mammography: improving detection and diagnosis of breast cancer," in *Computer Assisted Radiology and Surgery 2001, Proceedings of the 15th International Congress and Exhibition, Berlin, June 27-30, 2001.*, edited by H.U. Lemke, M.W. Vannier , K. Inamura, A.G. Farman, and K. Doi (Elsevier, Amsterdam: International Congress Series 1230), pp. 506-511.

Goodsitt M.M., H.P. Chan, J.T. Lydick, C.R. Gandra, N.G. Chen, M.A. Helvie, J. Bailey, M.A. Roubidoux, C. Paramagul, C.E. Blane, B. Sahiner, N. Petrick. 2004. An observer study comparing spot imaging regions selected by radiologists and a computer for an automated stereo spot mammography technique. Med Phys 31: 1558-1567.

Hayes R., M. Michell, and H.B. Nunnerley. 1991. Evaluation of magnification and paddle compression techniques in the assessment of mammographic screening detected abnormalities. Clinical Radiology 44:158-160.

Maidment A.D.A, P.R. Bakic, and M. Albert. 2003. Effects of quantum noise and binocular summation on dose requirements in stereoradiography, Med Phys 30: 3061-3071.

Petrick, N., H.P. Chan, B. Sahiner, M.A. Helvie. 1999. Combined adaptive enhancement and region-growing segmentation of breast masses on digitized mammograms. Med Phys 26: 1642-1654.

Petrick, N., H.P. Chan, B. Sahiner, M.A. Helvie, S. Paquerault, L.M. Hadjiiski. 2002. Breast cancer detection: Evaluation of a mass detection algorithm for computer-aided diagnosis: Experience in 263 patients. Radiology 224: 217-224.

Assessment of breast lesions on stereoscopic and monoscopic digital specimen mammograms: an ROC study

Heang-Ping Chan*, Mitch M. Goodsitt, Mark A. Helvie, Lubomir M. Hadjiiski, Justin T. Lydick, Marilyn A. Roubidoux, Janet E. Bailey, Alexis Nees, Caroline E. Blane, Joseph Shen, Michelle Foster, Berkman Sahiner

Department of Radiology, University of Michigan, Ann Arbor, MI 48109-0030

ABSTRACT

An ROC study was conducted to evaluate the usefulness of assessing breast lesion characteristics with stereomammography. Stereoscopic image pairs of 158 breast biopsy tissue specimens were acquired with a GE digital mammography system. Two stereo image pairs were taken at 1.8X magnification geometry and at approximately orthogonal orientations for each specimen. Display software was developed for a high resolution MegaScan CRT monitor driven by a DOME stereo display board. The specimens contained either a mass, microcalcifications, both, or normal tissue. About 40% of the specimens were found to contain malignancy by pathological analysis. Five MQSA radiologists participated in the observer performance experiment. The two views of each specimen were read independently and were separated by a large number of other specimen images to reduce any effects of memorization. Each observer read 316 specimen images in a systematically randomized order. The observer first read the monoscopic image and entered his/her assessment in terms of the confidence ratings on the presence of microcalcifications and/or masses, margin clearance, BI-RADS assessment, and the likelihood of malignancy. The corresponding stereoscopic images were then displayed on the same monitor and were viewed through stereoscopic LCD glasses. The observer was free to change the ratings in every category after stereoscopic reading. The ratings of the observers were analyzed by ROC methodology. For the 5 MQSA radiologists, the average A_z value for estimation of the likelihood of malignancy of the lesions improved from 0.70 for monoscopic viewing to 0.72 ($p<0.05$) after stereoscopic viewing, and the average A_z value for the presence of microcalcifications improved from 0.94 to 0.95 ($p<0.05$). The A_z value for the presence of masses improved from 0.80 to 0.82 after stereoscopic viewing, but the difference fell short of statistical significance ($p=0.08$). The visual assessment of margin clearance was found to have very low correlation with pathological analysis. This study demonstrates the potential of using stereomammography to improve the detection and characterization of mammographic lesions.

Keywords: Digital mammography, stereoscopic imaging, specimen radiograph, observer performance study.

1. INTRODUCTION

It is well known that mammographic sensitivity for breast cancer detection is limited by fibroglandular tissue overlapping with the lesions[1]. At least 25% of breast cancers in dense breasts are not identified[2, 3]. One main cause of these missed diagnoses can be attributed to the camouflaging effect of the overlapping structures. A number of new breast imaging techniques such as stereomammography[4-10], digital tomosynthesis[11, 12], and computed tomography[13, 14] are being developed in an effort to alleviate this problem.

We are developing stereomammography techniques using a full-field digital mammography system. Previously we have investigated the effects of stereo shift, geometric magnification, x-ray exposure, and display zooming on visual depth discrimination of crossing fibrils in stereo phantom images[5, 7, 15]. We have also studied the accuracy of using a calibrated virtual cursor to measure the absolute depth of fibrils in stereoscopic images[8-10]. In this study, we investigated whether stereoscopic viewing can improve the accuracy of detection and classification of mammographic lesions using breast tissue specimen images.

* chanhp@umich.edu

Medical Imaging 2004: Physics of Medical Imaging, edited by Martin J. Yaffe, Michael J. Flynn, Proceedings of SPIE Vol. 5368 (SPIE, Bellingham, WA, 2004) · 1605-7422/04/$15 · doi: 10.1117/12.535998

Stereoscopic imaging requires acquisition of a left-eye image and a right-eye image. Historically two film images were obtained by shifting the x-ray source, along a direction parallel to the image plane, to the left and the right of the central axis of the object. When the two film images are placed properly and viewed so that the left eye sees only the left-eye film and the right eye sees only the right-eye film, the parallax between the two images creates the depth perception. Stereoscopic imaging was utilized for various types of radiographic examinations. However, it did not achieve widespread acceptance in clinical practice, mainly because of the doubled film cost and increased patient exposure[16]. In the last few years, direct digital detectors have become available for medical imaging. Stereoscopic radiography may become a viable approach with digital imaging because no additional film costs will be required. Furthermore, digital detectors have wider dynamic range and higher contrast sensitivity than screen-film systems so that good-quality digital stereo image pairs may be acquired at the same or less total radiation dose than that for a conventional single-projection screen-film image[17]. Digital stereoscopic images can be viewed more conveniently than stereo film radiographs because of the electronic display. A stereoscopic LCD goggle is used as an electronic shutter. The goggle is synchronized with the display so that the left eye and the right eye of the reader are allowed to see only the left-eye image and the right-eye image, respectively, when the two images are displayed alternately at very fast refresh rate. This eliminates the need for a somewhat bulky film stereoscope or the requirement that the reader be trained to read stereoscopic images without aid.

An advantage of stereoscopic radiography is the three-dimensional (3D) structural information that it provides for the lesions of interest. The 3D distribution of microcalcifications may be associated with the malignant or benign nature of the cluster[18, 19]. The margin characteristics and spiculations of a mass may also be more readily distinguished from overlapping tissues under stereoscopic viewing conditions. The additional diagnostic information may improve the classification of malignant and benign lesions, thereby reducing unnecessary biopsies. It has been reported that digital stereomammography allowed the detection of additional lesions that were obscured on screen-film mammograms[4].

In this paper, we discuss the results of an observer performance study using receiver operating characteristic (ROC) methodology for the effects of stereo viewing on the detection and classification of breast lesions in biopsied tissue samples.

2. MATERIALS AND METHODS

2.1 Image Acquisition

Digital stereoscopic specimen image pairs were acquired with a GE Senographe 2000D digital mammography system. The system employs a digital detector consisting of a CsI:Tl scintillator and an amorphous-Si active matrix flat panel. The detector measures 23 cm x 18 cm, with a pixel size of 100 μm x 100 μm. We acquired stereoscopic pairs of images using a ±3° stereo-angle and 1.8X magnification (no grid, 0.15 mm focal spot) geometry. The target/filter, kVp and mAs were manually chosen. The target/filter combinations were mainly Mo/Mo with Mo/Rh in some cases. The mAs was chosen to be relatively high so that the images would not appear to be noisy. The images were acquired with a phantom-shift technique instead of a focal-spot shift technique. The technique has been discussed elsewhere[8, 15].

Biopsied breast tissue samples were imaged with the stereoscopic technique and 158 samples were used in the observer experiment. Stereo image pairs were acquired of the samples in two approximately orthogonal orientations by rolling the sample over by approximately 90°. The samples were randomly obtained without selection so that they can contain microcalcifications, mass, both, or normal tissue. Some samples were rejected because of improper shift between the left-eye and right-eye images. Otherwise all image pairs with good stereoscopic quality were included. About 40% of the samples were found to contain malignancy by pathological analysis.

2.2 Image Display

The images were displayed on a MegaScan 8M-pixel monitor driven by a Dome Md8-4820-LS stereoscopic board and a PC. The system can display full-field (2300 x 1800 pixels) digital mammograms at a refresh rate of 120 Hz.

It operates in a page flipping stereoscopic mode with the left- and right-eye images displayed alternately. A pair of CrystalEyes LCD stereoscopic glasses was used for viewing the stereoscopic images.

2.3 Observer Experiment

Five MQSA-qualified radiologists participated in the experiment. Each observer read 316 specimen images in a systematically randomized order. The two views of each specimen were read independently and were separated by a large number of other specimen images to reduce any effects of memorization. Prior to reading the test cases, the observers participated in training sessions to become familiar with the reading task. The observer first read the left-eye alone image as the monoscopic image and entered his/her assessment in terms of confidence ratings on the presence of microcalcifications and/or masses, margin clearance, BI-RADS assessment, and the likelihood of malignancy. The stereoscopic images were then displayed on the same monitor and were read with the LCD glasses. The observer was free to change his/her ratings in every category after stereoscopic reading. The ratings of the observers were analyzed with the LABMRMC program[20]. For the classification of malignant and benign lesions, all samples were analyzed together. The correlation of the readers' assessment of margin clearance with pathological analysis was quantified by the correlation coefficient.

3. RESULTS

Table 1 summarizes the average area under the ROC curves for the five MQSA radiologists. For the detection of microcalcifications, the average A_z value improved from 0.94 for monoscopic viewing alone to 0.95 with stereoscopic viewing. Although the standard deviations of the individual readers' A_z varied between 0.01 to 0.02, the improvement in A_z was consistent over all five radiologists, resulting in a statistically significant improvement ($p<0.05$). Similar result was observed for the classification of malignant and benign lesions in the tissue samples. The average A_z improved by 0.02 with stereoscopic viewing and the improvement was statistically significant ($p<0.05$). For the detection of masses, all five radiologists obtained an improvement in their individual A_z. However, their improvement in the average A_z value fell short of statistical significance ($p=0.08$). The assessment of margin clearance visually in tissue specimens was found to be very unreliable. The correlation coefficients of the margin clearance assessment by all radiologists with pathological analysis were all below about 0.3.

Table 1. Comparison of reading results for assessment of mammographic lesions in specimen radiographs without and with stereoscopic viewing.

Task	Average area under ROC curve (A_z)	
	Without Stereo Viewing	With Stereo Viewing
Detection of microcalcifications	0.94	0.95
Detection of masses	0.80	0.82
Classification of microcalcifications and masses	0.70	0.72
Margin clearance (Correlation Coefficients)	0.14-0.30	0.03-0.32

4. CONCLUSIONS

The purpose of this observer performance study was to evaluate the potential advantages of stereo full-field digital mammography for the detection and characterization of breast lesions. Detection and characterization of lesions in tissue samples is different from similar tasks in full-field digital mammograms. Since the location of the lesions are confined to a smaller and thinner tissue sample, specimen radiographs already provide superior visualization of lesion characteristics as compared to conventional mammography. Hence, it is more difficult to obtain further improvement in the detection and characterization tasks in specimen radiography. Nevertheless, our results indicate that the additional stereoscopic viewing did improve the visualization of lesions and the accuracy of assessing their malignant or benign characteristics in specimen images. It is likely that greater degrees of improvement in accuracy could be obtained in stereoscoping viewing of conventional mammograms. Although the results of this study cannot be generalized directly to reading full-field mammograms, the potential for information gain was demonstrated.

One limitation of stereoscopic viewing is that human eyes vary in their stereo acuity, although it is believed that stereo acuity may improve with training. Stereoscopic viewing also causes eye fatigue during prolonged reading, which may be alleviated with different display and viewing electronics and improved reader ergonomic issues.

ACKNOWLEDGMENTS

This work is supported by U. S. Army Medical Research and Materiel Command Grants DAMD 17-98-1-8210 and DAMD17-99-1-9294. The content of this publication does not necessarily reflect the position of the funding agency, and no official endorsement of any equipment and product of any companies mentioned in this publication should be inferred. The authors are grateful to Dr. Charles Metz for the LABMRMC program.

REFERENCES

1. V. P. Jackson, R. E. Hendrick, S. A. Feig and D. B. Kopans, "Imaging of the radiographically dense breast," *Radiology* 188, 297-301, 1993.
2. R. E. Bird, T. W. Wallace and B. C. Yankaskas, "Analysis of cancers missed at screening mammography," *Radiology* 184, 613-617, 1992.
3. M. G. Wallis, M. T. Walsh and J. R. Lee, "A review of false negative mammography in a symptomatic population," *Clinical Radiology* 44, 13-15, 1991.
4. D. J. Getty, R. M. Pickett and C. J. D'orsi, "Stereoscopic digital mammography: Improving detection and diagnosis of breast cancer," *Computer Assisted Radiology and Surgery 2001, Proc. 15th International Congress and Exhibition*, H. U. Lemke, M. W. Vannier, K. Inamura, et al., Eds. International Congress Series 1230, 506-511, Elsevier, Amsterdam, 2001.
5. H. P. Chan, M. M. Goodsitt, K. L. Darner, J. M. Sullivan, L. M. Hadjiiski, N. Petrick and B. Sahiner, "Effects of stereoscopic imaging technique on depth discrimination," *IWDM 2000- 5th International Workshop on Digital Mammography*, M. J. Yaffe, Ed., 13-18, Medical Physics Publishing, Madison, WI, 2001.
6. H. P. Chan, M. M. Goodsitt, L. Hadjiiski, M. A. Roubidoux, J. E. Bailey, M. A. Helvie, J. T. Lydick and B. Sahiner, "ROC study comparing radiologists' performances in evaluating breast lesions on stereoscopic and single-projection digital specimen mammograms," *Medical Physics* 30, 1456 (abstract), 2003.
7. H. P. Chan, M. M. Goodsitt, J. M. Sullivan, K. L. Darner and L. M. Hadjiiski, "Depth perception in digital stereoscopic mammography," *Program Book - Era of Hope Meeting, U. S. Army Medical Research and Materiel Command, Department of Defense, Breast Cancer Research Program.* 2000.
8. M. M. Goodsitt, H. P. Chan, K. L. Darner and L. M. Hadjiiski, "The effects of stereo shift angle, geometric magnification and display zoom on depth measurements in digital stereomammography," *Medical Physics* 29, 2725-2734, 2002.
9. M. M. Goodsitt, H. P. Chan and L. M. Hadjiiski, "Stereomammography: Evaluation of depth perception using a virtual 3d cursor," *Medical Physics* 27, 1305-1310, 2000.

10. M. M. Goodsitt, H. P. Chan, J. M. Sullivan, K. L. Darner and L. M. Hadjiiski, "Evaluation of the effect of virtual cursor shape on depth measurements in digital stereomammograms," M. J. Yaffe, Ed. Digital Mammography, 45-50, Medical Physics Publishing, Madison, WI, 2001.
11. S. Suryanarayanan, A. Karellas, S. Vedantham, S. P. Baker, S. J. Glick, C. J. D'orsi and R. L. Webber, "Evaluation of linear and nonlinear tomosynthetic reconstruction methods in digital mammography," *Academic Radiology* 8, 219-224, 2001.
12. L. T. Niklason, B. T. Christian, L. E. Niklason, D. B. Kopans, D. E. Castleberry, B. H. Opsahl-Ong, C. E. Landberg, P. J. Slanetz and E. Al., "Digital tomosynthesis in breast imaging," *Radiology* 205, 399-406, 1997.
13. J. M. Boone, T. R. Nelson and J. A. Seibert, "The potential for breast CT," *Medical Physics* 28, 1246 (abstract), 2001.
14. J. M. Boone, T. R. Nelson, K. K. Lindfors and J. A. Seibert, "Dedicated breast CT: Radiation dose and image quality evaluation," *Radiology* 221, 657-667, 2001.
15. H. P. Chan, M. M. Goodsitt, L. M. Hadjiiski, J. E. Bailey, K. Klein, K. L. Darner and B. Sahiner, "Effects of magnification and zooming on depth perception in digital stereomammography: An observer performance study," *Physics in Medicine and Biology* 48, 3721 - 3734, 2003.
16. T. S. Curry, J. E. Dowdey and R. C. Murry, *Christensen's physics of diagnostic radiology.*, 4th ed. Pages, Lea & Febiger, 1992.
17. A. D. A. Maidment, P. R. Bakic and M. Albert, "Effects of quantum noise and binocular summation on dose requirements in stereoradiography," *Medical Physics* 30, 3061-3071, 2003.
18. A. D. A. Maidment, M. Albert, E. F. Conant, C. W. Piccoli and P. A. Mccue, "Prototype workstation for 3-d diagnosis of breast calcifications," *Radiology* 201(P), 556, 1996.
19. E. F. Conant, A. D. Maidment, M. Albert, C. W. Piccoli, S. A. Nussbaum and P. A. Mccue, "Small field-of-view digital imaging of breast calcifications: Method to improve diagnostic specificity," *Radiology* 201(P), 369, 1996.
20. D. D. Dorfman, K. S. Berbaum and C. E. Metz, "Roc rating analysis: Generalization to the population of readers and cases with the jackknife method," *Investigative Radiology* 27, 723-731, 1992.